Development and Use of Genetic Reporters to Identify *Mycobacterium tuberculosis* Exported Proteins Important for Virulence

Jessica Rose McCann

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the department of Microbiology and Immunology, School of Medicine

Chapel Hill
2009

Approved by
Advisor: Miriam Braunstein
Reader: Robert Bourret
Reader: Janne Cannon
Reader: Jeffrey Dangl
Reader: Thomas Kawula

ABSTRACT

Development and use of genetic reporters to identify *Mycobacterium tuberculosis* exported proteins important for virulence

Jessica Rose McCann
(Under the direction of Miriam Braunstein)

Mycobacterium tuberculosis is the causative agent of tuberculosis, a disease that kills nearly two million people each year globally. To cause disease, *M. tuberculosis* must survive and replicate in macrophages in the lung within phagosomal compartments. Macrophages typically neutralize bacterial invaders, but *M. tuberculosis* has developed defenses against macrophage attack. The relatively recent ability to construct *M. tuberculosis* mutants has greatly advanced our understanding of the protein transport systems and the importance of exported proteins to virulence of this pathogen. Exported proteins, which we define here as those exposed on the surface or exported out of the bacterium, make excellent bacterial virulence factors as they are best localized to counteract host defenses against infection. The main objective of the research described in this dissertation was to develop genetic tools and then use them to identify *M. tuberculosis* exported proteins important to pathogenesis. For many of the proteins identified in this research, these studies provide the first experimental evidence of export. In the chapters that follow, three different genetic export reporter systems are described. We constructed a new reporter system to specifically identify *M. tuberculosis* proteins exported by the twin arginine translocation pathway. We also developed a new reporter system with the unique ability to report on protein export during *M. tuberculosis* growth in host cells. This system was then modified into

a new reporter transposon used to simultaneously identify exported proteins of *M. tuberculosis* and collect mutants lacking these proteins. The resulting library of *M. tuberculosis* mutants was successfully screened for mutants defective for growth in macrophages as a strategy to identify new exported virulence factors. Future investigations are sure to focus more directly on how proteins exported by this pathogen interact with the host and cause disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

2D	two-dimensional
A_{486}	absorbance at 486 nanometers
ABC	adenosine tri-phosphate binding cassette
AIDS	Acquired Immune Deficiency System
aph	kanamycin resistance gene
ATP	adenosine tri-phosphate
BCA	bisinchoninic acid
BCG	bacillus Calmette-Guérin
Bla	beta-lactamase
bp	base pair
C-terminus	carboxy-terminus
carb	carbenicillin
CFU	colony-forming unit
DNA	deoxyribnucleic acid
E.	*Escherichia*
ESAT-6	early secreted antigen target 6 kilodalton
ESX-1	ESAT-6 secretion system
FCS	fetal calf serum
HA	hemagglutinin
HIV	Human Immunodeficiency Virus
hyg	hygromycin resistance gene
K	lysine
kbp	kilobase pair

kDa	kilodalton
L	leucine
M	methionine
M.	*Mycobacterium*
Mbp	megabase pair
MDR	multi-drug resistant
mg	milligram
MIC	minimum inhibitory concentration
min	minute
ml	milliliter
MOI	multiplicity of infection
N-terminus	amino-terminus
ng	nanogram
nm	nanometer
OD600	optical density, 600 nanometers
ORF	open reading frame
ori	origin of replication
P.	*Pseudomonas*
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Plc	phospholipase C
PMA	phorbol myristate acetate
R	arginine
RNA	ribonucleic acid

SCID	Severe Combined Immunodeficiency
SDS	sodium dodecyl sulfate
Sec	general secretion pathway
SRP	signal recognition particle
ss	signal sequence
Tat	twin-arginine translocation pathway
TB	tuberculosis
TraSH	transposon site hybridization
WCL	whole cell lysate
XDR	extensively drug-resistant
β	beta
Δ	deletion
μg	microgram
μl	microliter
μm	micrometer
::	gene insertion

CHAPTER 1

INTRODUCTION

Mycobacterium tuberculosis was discovered as the causative agent of tuberculosis (TB) over 125 years ago. Even so, the virulence mechanisms used by this successful bacterial pathogen remain poorly understood. TB kills nearly two million people a year and it is estimated that a third of the world's population is currently infected with *M. tuberculosis* (105). Most individuals who become infected have a one in ten chance of developing active TB over the course of their lifetime. However, that risk of disease increases to 10% each year for people with HIV. In fact, one of the major forces behind the increase in tuberculosis disease is the deadly synergy between HIV and TB. A significant proportion of people who die of AIDS actually succumb to *M. tuberculosis* infections (28). Another emerging problem in TB disease is the rise of multiply and extensively drug resistant strains of *M. tuberculosis* (MDR and XDR, respectively) (106).

TB is spread from person to person when an actively infected individual coughs and sprays aerosolized bacteria into the air, which is then inhaled by a passerby (34, 86). If the bacteria travel deep into the alveoli of the lung, they are engulfed by macrophages. Macrophages typically phagocytose foreign organisms and disassemble them into presentable antigens. They thereby serve as alert sentinels of the immune system. *M. tuberculosis*, however, is an intracellular pathogen able to avoid this destruction and grow in macrophages. A hallmark of intracellular *M. tuberculosis* is that it resides in phagosomes that do not mature normally into acidified phagolysomes. The ability of the bacterium to alter phagosome maturation is one activity that may be important for intracellular survival (43, 103). *M. tuberculosis* survival in macrophages may also depend on the ability of the bacterium to resist attack of reactive radicals, inhibit host immune responses and/or manipulate programmed cell death pathways of

host cells (33, 43, 73, 83). Later in *M. tuberculosis* infection, most immunocompetent individuals produce an antigen specific T cell response that activates macrophages to become more antimycobacterial and establishes granulomas to control growth of the bacteria and the initial infection. However, another impressive feature of *M. tuberculosis* is its ability to persist long-term in granulomas and later reactivate to cause active TB disease. The many aspects of disease caused by *M. tuberculosis* can be studied in mice and cultured macrophages. These model systems are useful tools for studying the virulence factors *M. tuberculosis* employs to survive and spread in the host.

Numerous efforts are underway to identify the *M. tuberculosis* virulence factors involved in the above activities. The exported proteins of bacterial pathogens, which we will define as proteins transported beyond the cytoplasm to the bacterial cell surface or secreted into the environment, are well-positioned to modify the host and/or protect the bacterium from antimicrobial attack. This is especially true for intracellular pathogens, whose contact with the host is as intimate as can be achieved. There are now several examples of *M. tuberculosis* exported proteins that are proven virulence factors (48). One such exported virulence factor is *e*xported *r*epetitive *p*rotein, or Erp (5, 6). However, the function of Erp in virulence remains unknown.

Mycobacteria possess a complex and unique cell envelope (11, 23) (Fig. 1.1). Given that the mycobacterial cell wall constitutes a formidable barrier to drugs and the host immune system, it is also likely to pose a challenge to directed export of proteins out of the bacterial cytosol. At the core of the mycobacterial cell envelope is a covalently linked network of peptidoglycan, arabinogalactan and novel long chain mycolic acids. The complex lipids phthiocerol dimycocerosate (PDIM) and trehalose dimycolate (TDM) are associated with the core. PDIM and TDM are non-covalently attached to the mycolic acid layer. Exported proteins can be found integrated into the cytoplasmic membrane, peripherally associated with the cytoplasmic membrane, distributed throughout the cell wall layers, or secreted out into the extracellular space. As may be inferred from its complex cell envelope structure,

M. tuberculosis is not a typical Gram-negative or Gram-positive bacterium. Sometimes identified as Gram-positive, *M. tuberculosis* does not Gram stain well and instead is identified by acid-fast staining.

The objective of this dissertation was to develop genetic tools that could then be used to identify *M. tuberculosis* exported proteins important for pathogenesis. In this introductory Chapter, I will first describe what is known about the conserved and specialized protein export pathways in *M. tuberculosis*. I will then discuss other genetic systems used previously to identify *M. tuberculosis* exported proteins. Finally, I will review genome-wide strategies that have been used to identify *M. tuberculosis* mutants attenuated for growth in macrophage and mouse models of tuberculosis.

Conserved Mycobacterial Export Pathways

Sec Pathway

The Sec pathway is a conserved protein export system found in all bacteria (67, 74). This pathway recognizes proteins that are synthesized as precursors with a characteristic N-terminal tripartite signal sequence. It transports these proteins in an unfolded state from the cytosol across the cytoplasmic membrane to the cell envelope. Once the protein has traversed the cytoplasmic membrane, its signal sequence is cleaved and the protein folds into its final conformation. The signal sequence is critical for recognition by the Sec pathway, and the conserved elements of these signal sequences enable predictions of Sec exported proteins using bioinformatic algorithms (29). Proteins exported out of the cytosol by the Sec pathway can permanently reside in the cell envelope or be further secreted into the environment through the action of subsequent systems.

At the core of the Sec pathway is a membrane spanning translocase channel, composed of the integral membrane SecY, SecE and SecG proteins (67, 74). Another central component is the SecA ATPase which provides energy for translocation. The Sec pathway is the major protein export system of bacteria. It is responsible for the export of numerous substrates, many of which have essential roles in cell physiology. Thus, it is not surprising that the Sec pathway is essential for viability in all bacteria in

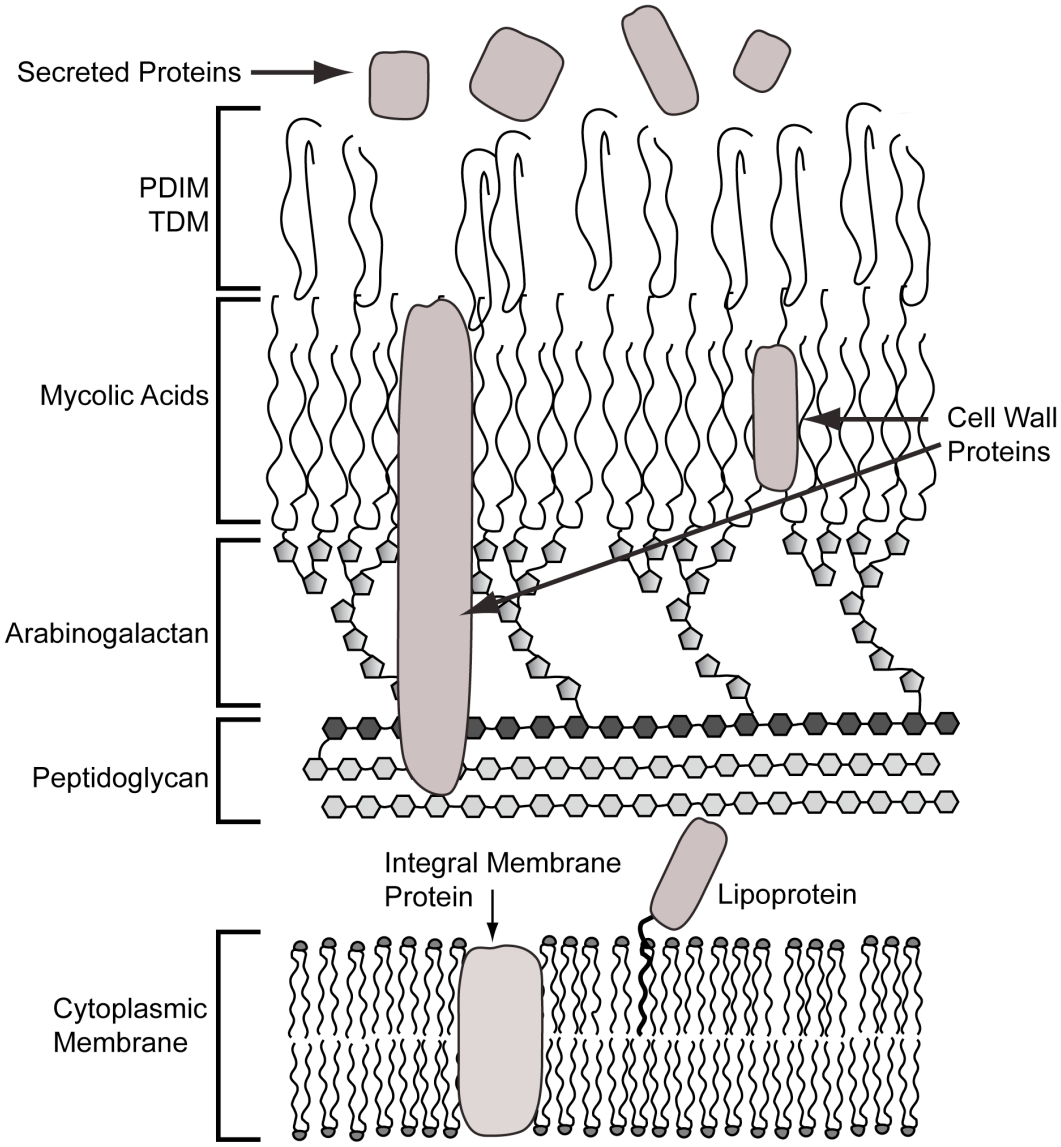

Figure 1.1. A model of the *M. tuberculosis* cell envelope. *M. tuberculosis* has a complex cell envelope that is considered neither Gram-negative nor Gram-positive. PDIM, phthiocerol dimycocerosate. TDM, trehalose dimycocerosate. Different classes of exported proteins are indicated.

which it has been tested. The Sec pathway also exports many examples of bacterial virulence factors (30). Mycobacteria have a functional Sec pathway. Orthologs of all the essential Sec components are present in mycobacterial genomes, including pathogenic *M. tuberculosis* and non-pathogenic *M. smegmatis* (7, 21). Further support for a functional mycobacterial Sec pathway comes from proven surface and secreted proteins of *M. tuberculosis* that are synthesized with Sec signal sequences (48). For some of these proteins there is also more direct evidence of signal sequence cleavage (61).

Interestingly, all mycobacterial genomes analyzed encode two SecA homologs: SecA1 and SecA2 (8). This is highly unusual, although it is now recognized that there are also some Gram-positive bacteria with two SecAs (4, 18, 55). Both mycobacterial SecAs are highly homologous to the single well-studied SecA of *E. coli* and both have demonstrated ATPase activity (42). However, SecA1 has an overall higher level of sequence similarity to *E. coli* SecA. SecA1 is also the essential SecA in *M. smegmatis* and *M. tuberculosis,* like the single SecA of other bacteria, while SecA2 is not (8). Finally, experiments using a conditional *secA1 M. smegmatis* strain also show SecA1 to be necessary for export of the Sec signal sequence-containing MspA porin to the cell wall (40). On the basis of these results, SecA1 is believed to be the "housekeeping" SecA. In contrast, SecA2 is believed to be a "specialized" SecA. The SecA2-dependent export pathway is discussed in detail below.

Tat Pathway

The twin-arginine translocation (Tat) pathway is another conserved protein export pathway (54). This pathway is present in many, but not all, bacteria (27). Like the Sec pathway, the Tat pathway exports proteins synthesized with N-terminal tripartite signal sequences across the cytoplasmic membrane. Following protein export the signal sequence is cleaved, and the proteins may remain associated with the cell envelope or be further secreted (54, 104). Unlike Sec-exported proteins, Tat substrates typically have a twin-arginine motif in the charged domain of their signal sequence designated R-R-x-ϕ-ϕ (ϕ=hydrophobic). Site-directed mutagenesis experiments involving conservative substitution

of the 'RR' dipeptide with 'KK' demonstrate the importance of the twin-arginine motif (97). Other Tat targeting elements also exist (102). Another difference between the Sec and Tat pathways is that the Tat pathway exports proteins in a folded state and relies on energy from the proton motive force.

The Tat translocation apparatus is comprised of the integral membrane proteins TatA, TatB, and TatC. Current models hypothesize that TatA forms a homo-oligomeric complex that serves as the channel through which the protein is transported and that TatB and TatC are important for recognition of the precursor protein and delivery to the channel (54).

Orthologs of TatA, TatB, and TatC are present in mycobacteria. Tat signal sequence prediction programs also identify candidate Tat-exported substrates (27). Mutational analysis demonstrates the Tat pathway is functional in mycobacteria. In the fast growing and non-pathogenic *M. smegmatis,* mutants deleted for *tatA, tatB,* or *tatC* were constructed (65, 76). These *tat* mutants have a growth defect on nutrient agar as well as several other phenotypes including hypersensitivity to the detergent sodium dodecyl sulfate and sensitivity to β-lactam antibiotics. The pleiotropic effects of the *tat* mutations imply there are multiple Tat exported proteins. They include the mycobacterial β-lactamases, which possess functional Tat signal sequences and require export by the Tat pathway to protect the bacterium from β-lactam antibiotics. In *M. tuberculosis* it has not been possible to delete *tatA, tatB,* or *tatC* (87). Thus, at least under standard laboratory conditions, the Tat pathway is essential in slow-growing virulent mycobacteria.

There are several examples of bacterial pathogens in which the Tat pathway contributes to virulence (15, 25, 53, 68, 77). One of the best characterized examples of this is the pathogen *Pseuodomonas aeruginosa. P. aeruginosa tat* mutants are attenuated in an animal model of virulence. The phospholipase C (Plc) enzymes, which are proven *P. aeruginosa* virulence factors, are synthesized with Tat signal sequences and exported in a Tat-dependent manner (68, 95).

There are also Plc enzymes with Tat signal sequences and demonstrated roles in virulence in *M. tuberculosis* (45, 79). For PlcB of *M. tuberculosis*, the N-terminal signal sequence was shown to be

functional in promoting Tat export using a β-lactamase reporter system (65). In this reporter system, a truncated *M. tuberculosis* β-lactamase ('BlaC) lacking a signal for export is fused to a foreign signal sequence and export of the resulting 'BlaC fusion protein is demonstrated by the ability of the reporter to protect against cell wall targeting β-lactam antibiotics. When expressed in *M. smegmatis,* a PlcB signal sequence fused to 'BlaC is able to protect from β-lactam as long as a functional Tat pathway is present. In addition, in both *M. smegmatis* and *M. tuberculosis* conservative substitution of the twin 'RR' dipeptide with a 'KK' dipeptide in the PlcB signal sequence of the fusion eliminates export (65). These experiments demonstrate that the *M. tuberculosis* PlcB signal sequence is functional in Tat export, and it suggests that the Tat pathway is important to the virulence of *M. tuberculosis*. The Tat pathway and its substrates in *M. tuberculosis* are further discussed in Chapter 2 of this dissertation.

Specialized Mycobacterial Export Pathways

In numerous Gram-negative bacterial pathogens, there exist specialized protein export pathways that secrete virulence factors across the bacterial outer membrane and often directly translocate them into host cells (19, 75, 88). Specialized export systems also exist for Gram-positive pathogens and mycobacteria, although less is known about these more recently identified systems.

SecA2-Dependent Protein Export Pathway

As mentioned above, an unusual property of mycobacteria is the presence of two SecAs. Although it is not a universal property of Gram-positive bacteria, there are some examples of Gram-positive organisms with two SecAs. These examples include several pathogens: *Listeria monocytogenes*, *Streptococcus gordonii*, and *Streptococcus parasanguis* (4, 18, 55). As discussed above, SecA1 is believed to be the "housekeeping" SecA with the major function in transporting proteins beyond the cytosol. On the other hand, SecA2 is not essential and is considered the

"specialized" SecA protein. The *secA2* mutant of *M. smegmatis* has several altered phenotypes when compared to wild-type organisms (8). The *secA2* mutant of *M. tuberculosis* is attenuated for growth in cultured macrophages and virulence in mice, which demonstrates a role for SecA2 in pathogenesis (49). Like SecA1, SecA2 has proven ATPase activity (42). Further, a Δ*secA2* mutant of *M. smegmatis* or *M. tuberculosis* cannot be complemented with SecA2 bearing a conservative amino acid substitution in its ATP binding domain indicating that ATPase activity is required for its function (42, 84). Using *M. smegmatis* it was shown that the SecA1 and SecA2 proteins are not functionally redundant (8). Mutant phenotypes associated with the absence of one of the proteins cannot be suppressed by overexpression of the other, indicating that SecA1 and SecA2 have different roles in mycobacteria.

Substrates of the SecA2 Export Pathway

Experiments with the *secA2* mutants of *M. smegmatis* and *M. tuberculosis* demonstrate a role for SecA2 in protein export and, more specifically, in the export of a specific subset of proteins. In *M. smegmatis* (37) a proteomic approach identified cell envelope proteins that depend on SecA2 for localization. Comparison of the protein profiles of cell wall and cell membrane fractions from the *secA2* mutant and wild-type *M. smegmatis* using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) revealed the majority of proteins were unaffected by SecA2. However, two proteins (Msmeg 1704 and Msmeg 1712) are exported to these fractions only when SecA2 is present. Both these SecA2-dependent proteins are lipoproteins with sequence similarity to periplasmic sugar binding proteins of other bacteria, and both have predicted Sec signal sequences. It is important to note that not all lipoproteins in *M. smegmatis* are SecA2-dependent. In fact, the majority of *M. smegmatis* lipoproteins are unaffected by SecA2 indicating specificity in the substrates recognized by SecA2.

A similar proteomics study in *M. tuberculosis* (9) compared the proteins secreted into the culture filtrate by *secA2* mutant and wild-type *M. tuberculosis* using 2D-PAGE. The culture filtrate is a cell-free fraction composed of the proteins secreted into the culture medium during exponential growth. Three proteins show diminished export in a *secA2* mutant, including the antioxidant enzymes superoxide

dismutase (SodA) and catalase-peroxidase (KatG). In contrast to the SecA2 dependent lipoproteins found in *M. smegmatis*, SodA and KatG of *M. tuberculosis* are secreted via SecA2 but lack any recognizable signal sequence (9). Interestingly, *L. monocytogenes* also exports a superoxide dismutase (MnSod) lacking a signal peptide in a SecA2 dependent manner (2).

The discovery that antioxidants are secreted via SecA2 has interesting implications for understanding how *M. tuberculosis* SecA2 contributes to virulence. This is because bacterial antioxidants secreted into the host could protect bacteria from the harmful effects of the oxidative burst of macrophages. This could be part of the explanation for the role of *M. tuberculosis* SecA2 in virulence. However, *M. tuberculosis* SecA2 also limits host immune responses and promotes growth in macrophages even in the absence of the oxidative burst (50). Therefore, protection from oxygen radicals can only be part of the function of SecA2 in virulence.

How SecA2 in mycobacteria works remains to be elucidated. Because there are both proteins with and proteins without signal sequences that have been identified as SecA2-dependent, there may in fact be multiple roles for SecA2 in promoting protein export. One role may be in the export of a subset of signal sequence-containing proteins, such as those identified in *M. smegmatis*. Another role may be in the export of substrates lacking a conventional signal sequence, such as those identified in *M. tuberculosis* (Fig. 1.2). There is evidence that for the *M. smegmatis* SecA2-dependent proteins containing signal sequences SecA2 works in conjunction with SecA1. Depletion of SecA1 in *M. smegmatis*, reduces export of the SecA2 dependent substrate Msmeg1712 to a level comparable that observed in a *secA2* mutant (84). It is possible that SecA2 is required for the delivery of certain substrates to the canonical SecA1/SecYEG pathway, where export then continues in a conventional manner. For the examples of non-signal sequence bearing SecA2-dependent proteins, SecA2 could directly participate in the transport of these proteins either working in conjunction with the conserved SecA1/SecYEG channel or with an unknown channel. Another possibility is that the role of SecA2 in the export of proteins without signal sequences is indirect. SecA2 could actually export a signal

sequence-containing protein to the cell envelope which is itself a component of a novel secretion system for unconventional exported proteins such as SodA (Fig. 1.2).

ESX-1 secretion pathway

A second specialized export system initially identified in mycobacteria and now recognized to also exist in some Gram-positive bacteria is the ESAT-6 secretion system (Esx-1) (14, 35, 72). This system is named for the small highly immunogenic protein known as the Early Secreted Antigenic Target 6 (ESAT-6, EsxA, Rv3875) that is secreted by *M. tuberculosis* via this pathway (1, 96). In complex with ESAT-6, the ESX-1 system secretes an ESAT-6 like protein named Culture Filtrate Protein 10 (CFP-10, EsxB, Rv3874). In addition, there are two more secreted substrates of this system which are not ESAT-6 family members, EspA (Rv3616c) and EspB (Rv3881c), which are discussed further below (32, 66). Both ESAT-6 and CFP-10 are well-established as being secreted proteins but both lack conventional Sec or Tat signal sequences. ESAT-6 and CFP-10 are encoded by adjacent genes in an operon located in a region of the *M. tuberculosis* genome termed RD1 (for *R*egion of *D*eletion). RD1 is absent from the attenuated *M. bovis* BCG vaccine strain which explains why these proteins are absent from BCG (3, 41, 60).

It is now recognized that 23 ESAT-6 homologs (EsxA-W) are encoded by the *M. tuberculosis* genome. Each of these proteins is small, about 100 amino acids in length, and most have a WXG motif in the C-terminus (35, 72, 101). All ESX proteins lack a traditional signal sequence but many are detected in the culture filtrate. Thus, secretion is a characteristic of many, if not all, of the members of this protein family (17, 61, 64, 85, 94). Like *esxA (esat-6)* and *esxB (cfp-10)*, most of the *esx* genes in *M. tuberculosis* exist as tandem pairs in the genome. Five of the paired *esx* loci exhibit further similarity in being clustered with a suite of homologous genes. These five *esx* genomic regions have been designated ESX-1 through ESX-5 systems (35, 101), with ESX-1 (spanning open reading frames

Rv3866 to Rv3883c) being the first and best described of the systems and the system containing the genes for the prototype ESAT-6 (EsxA) and CFP-10 (EsxB) proteins (Fig. 1.3).

In specialized secretion systems of other bacterial pathogens it is common to find the secreted effectors encoded in the same genomic region as the secretion apparatus. Inspection of the other open reading frames (ORFs) in the ESX loci reveals conserved proteins with predicted properties that suggest functions as part of a secretion system. These include predicted nucleotide binding proteins, membrane proteins, and chaperones.

The ability to conduct genetic manipulation of *M. tuberculosis* and *M. bovis* BCG has defined the genes in ESX-1 required for ESAT-6 and CFP-10 secretion. Transposon insertions and deletions of individual ESX-1 genes in *M. tuberculosis* as well as expression of ESX-1 components in BCG, demonstrates that the many of the genes surrounding the *esxAB* operon in ESX-1 are necessary for secretion (13, 39, 44, 78, 98). With the same approach the ESX-1 system was also shown to be important for the virulence of *M. tuberculosis* (39, 56).

In all, *M. tuberculosis* gene products spanning Rv3860 to Rv3879c, which includes some ORFs beyond the boundaries of ESX-1, have been tested for their involvement in ESAT-6 secretion. Rv3870, Rv3871, and Rv3877 are all demonstrated to be required for ESAT-6 and CFP-10 secretion and can be considered core components of the ESX-1 secretion apparatus (13, 39, 44, 98). Rv3870 and Rv3871 are predicted subunits of a membrane-bound AAA ATPase of the FtsK/SpoIIIE family. Together these two proteins possess three ATP/GTP binding motifs (35), and protein interaction studies show that Rv3871 directly interacts with Rv3870 (98). These proteins are attractive candidates for providing energy necessary for export. Rv3871 further interacts with CFP-10 (93, 98), which links a secreted substrate to the secretion apparatus. Rv3877 has eleven predicted transmembrane domains (stretches of hydrophobic amino acids predicted to span the cytoplasmic membrane) making it reminiscent of the SecY translocase protein and best candidate so far for being a membrane-spanning channel. So far,

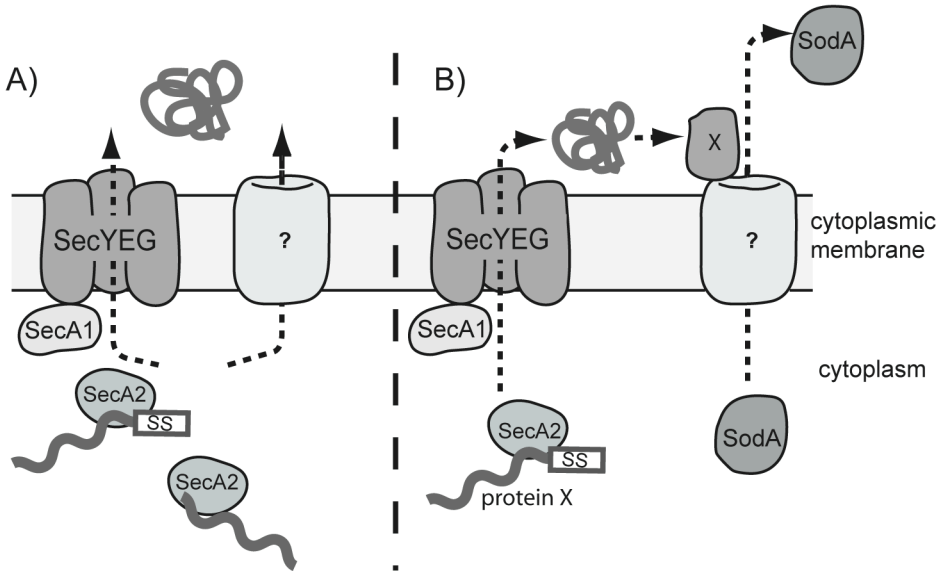

Figure 1.2. Potential models for SecA2-dependent export. SecA1 and SecYEG channel are involved in export of proteins with signal sequences, while the function of SecA2 remains unclear. **A)** SecA2 may directly export proteins with or without signal sequences through either the SecYEG channel or an unknown channel. **B)** SecA2 may indirectly export proteins without signal sequences by first exporting factors ("protein X") that function as part of a multiprotein complex/channel responsible for the export of proteins lacking conventional signal sequences such as SodA.

there are no physical interactions mapped between Rv3870/71 and Rv3877, but it is easy to imagine that these proteins work together. Rv3868, a demonstrated AAA ATPase (58), and Rv3869 proteins are also required elements of the ESX-1 secretion system (13). Rv3689 is a predicted membrane protein lacking any obvious functional domain (35). As in most of the ESX systems, a tandem pair of genes encoding PE and PPE proteins is present in ESX-1 (36). These protein families are named for the Pro-Glu (PE) or Pro-Pro-Glu (PPE) motifs present in conserved N-terminal domains. Although a large percentage (10%) of the *M. tuberculosis* genome encodes these proteins their function is unknown. Rv3872 (PE35) is necessary for ESAT-6/CFP-10 secretion while Rv3873 (PPE68) is not (13, 26, 44). PE35 may have a regulatory function, as *esxAB* genes are not transcribed in an *rv3872* mutant (13). Of note, PE35 and PPE68 themselves appear to be exported as they are detected in *M. tuberculosis* culture filtrate and cell envelope, respectively (26, 32, 70). Other highly conserved ORFs in ESX-1 that are worthy of mention are Rv3876 (predicted ATPase), Rv3879c (predicted cytosolic protein of unknown function), Rv3881c (EspB, secreted ESX-1 substrate), Rv3882c (predicted membrane protein), and Rv3883c (MycP1, predicted serine protease associated with the cell wall) (24). For Rv3876 and Rv3879c, directed mutation showed them not to be required for ESAT-6 and CFP-10 secretion while the other three ORFs remain untested in *M. tuberculosis* (13). Along with genes in ESX-1, it is now recognized that there are other ORFs elsewhere in the genome involved in ESAT-6/CFP-10 secretion. Rv3614c, Rv3615c, and Rv3616c (EspA) are needed for ESAT-6/CFP-10 secretion (32, 59). Interestingly, these three ORFs have homology to Rv3864, Rv3865 and Rv3867, which are all located at the edge of ESX-1. However, inactivation of the paralogous genes at the ESX-1 locus does not result in an ESAT-6/CFP-10 secretion defect (13). There is also a connection between Rv3614c and ESX-1 ORFs in that yeast two-hybrid analysis shows that Rv3614c interacts with Rv3882c as well as interacting with itself and Rv3615c (MacGurn *et al.*, 2005).

Not only is EspA required for ESAT-6 and CFP-10 secretion but EspA is itself a secreted ESX-1 substrate requiring ESAT-6 and CFFP-10 for its own secretion (32). EspA is a 40 kDa protein lacking a

traditional N-terminal signal sequence with no obvious homology to ESAT-6 family members. The co-dependence in secretion between ESAT-6, CFP-10, and EspA suggests that these three proteins interact prior to secretion or are somehow involved in the secretion process itself. Perhaps they stabilize each other during export or serve as structural components of the actual ESX-1 secretion apparatus (Fig. 1.3). Recently, Rv3881c (EspB) which is encoded in the ESX-1 locus was reported to be a fourth ESX-1 secreted substrate of *M. tuberculosis* (66). EspB is a 55 kDa protein rich in alanine and glycine residues that also has no obvious signal sequence for export and no homology to EspA or ESAT-6 proteins.

How the above described proteins work together to secrete ESAT-6, CFP-10, EspA, and EspB is actively under investigation. The identification of protein-protein interactions between ESX-1 components and secreted substrates is making the biggest impact on our understanding of the system.

ESAT-6 and CFP-10 physically interact to form a stable 1:1 heterodimer *in vitro* and ESAT-6/CFP-10 complexes have been identified prior to secretion (in the cytoplasm) and after secretion (in culture filtrates) of *M. tuberculosis* which indicates the proteins are secreted as a complex (12, 69, 81, 82, 93, 98). One function of this complex seems to be to stabilize the individual proteins (39, 82). Another function is in targeting the complex for secretion. The C-terminal seven amino acids of CFP-10 serve as a targeting sequence required for secretion of the heterodimer out of the cell (17). A truncated CFP-10 lacking these seven amino acids still binds to ESAT-6 but the complex is no longer secreted. Not only is the C-terminal signal necessary for secretion but it is sufficient to promote secretion by *M. tuberculosis* of ubiquitin, a small heterologous protein (17). Yeast two hybrid studies show that the C-terminal region of CFP-10 interacts with the predicted ATPase subunit Rv3871which could then provide the energy necessary for secretion of the ESAT-6/CFP-10 complex (17, 93, 98) (Fig. 1.3). Other interactions have been mapped for the ESX-1 system and they reveal the complexity of the system (Fig. 1.3). Western-western blotting shows that ESAT-6 and CFP-10 bind to PPE68 (69). As mentioned above, interactions are also reported between Rv3614c and Rv3615c and Rv3882c(59). Finally, EspB

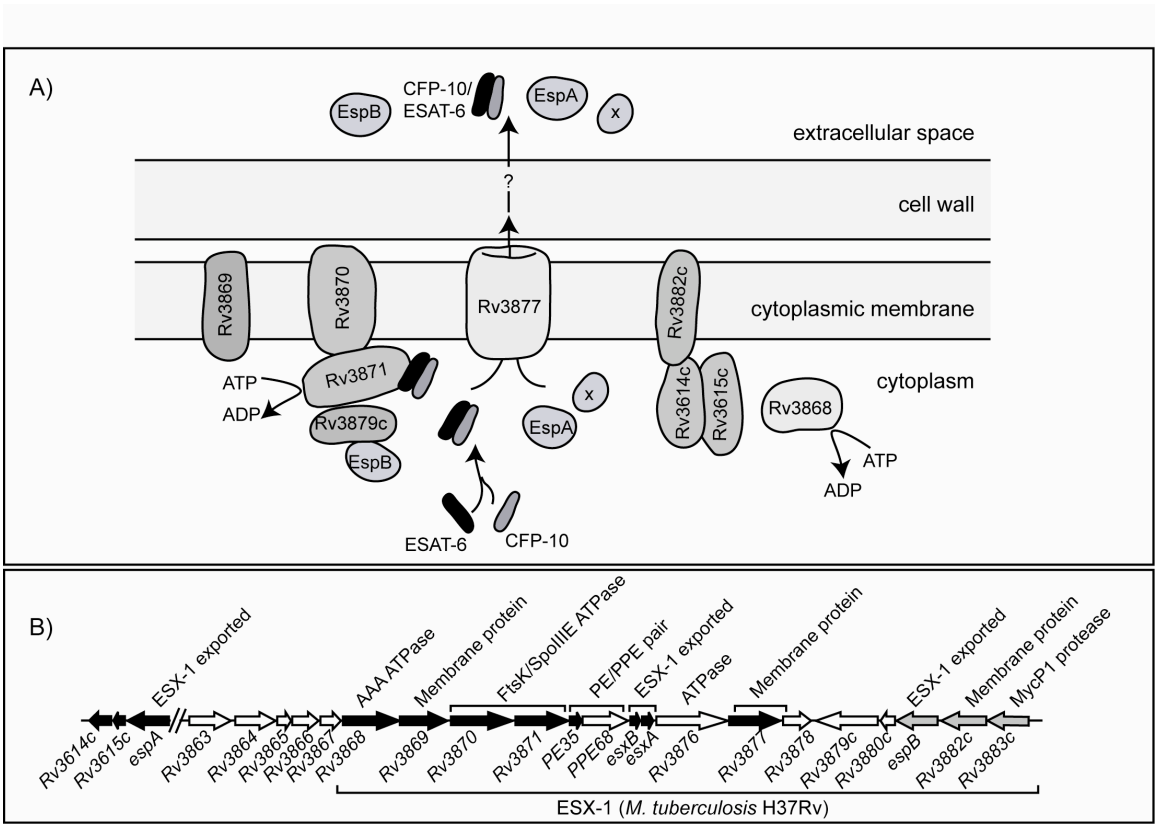

Figure 1.3 The ESX-1 secretion pathway of *M. tuberculosis*. **A)** Potential model of ESX-1 secretion in *M. tuberculosis*. ESAT-6, CFP-10, EspA, EspB and possibly other unknown proteins (depicted here as protein X) are secreted by the ESX-1 system made up of predicted membrane proteins, ATPases, and cytoplasmic chaperones. Rv3877 is the leading candidate for being a translocation channel spanning the cytoplasmic membrane. How the proteins are further secreted across the cell wall remains a mystery. **B)** Genetic organization of genes shown by mutation as being required for ESX-1 secretion in *M. tuberculosis*. Black ORFs encode proteins demonstrated as required for the secretion of ESAT-6 and CFP-10 in *M. tuberculosis*; white ORFs encode proteins not required for export in *M. tuberculosis*; gray ORFs are untested in *M. tuberculosis*. ESAT-6 and CFP-10 are encoded by *esxA* and *esxB*, respectively.

binds Rv3879c (66). Rv3879c in turn binds to Rv3871 (66). This last set of interactions is particularly interesting since EspB secretion by *M. tuberculosis* depends on ESX-1, but it is CFP-10-independent. This suggests that CFP-10 is not responsible for routing EspB to the ESX-1 secretion machine. The interactions suggest that *M. tuberculosis* Rv3879c, which is not required for ESAT-6 or CFP-10 secretion (13) but is untested for EspB secretion, could be the missing link. Even with the increasing examples of protein-protein interactions in the ESX-1 system many aspects of ESAT-6/CFP-10 secretion remain a mystery (Fig. 1.3).

A Genetic System for Identification of *M. tuberculosis* Exported Proteins

Approximately 20% of the *M. tuberculosis* genome is predicted to encode exported proteins based on scans using bioinformatic algorithms that identify signal sequences or transmembrane domains (21, 92). As mentioned previously, virulence factors of bacterial pathogens are often exported, so a high priority is placed on their identification. Further, exported proteins are often immunogenic and make good targets for diagnostic antigens. Prior efforts to identify exported proteins of *M. tuberculosis* have used biochemical and genetic methods (48). In this dissertation, I discuss large-scale genetic strategies we developed to identify *M. tuberculosis* exported proteins that use new genetic reporter constructs. Here, I will review a previous genetic strategy used to identify *M. tuberculosis* exported proteins and discuss its limitations.

A reporter for protein export is typically an enzyme that requires translocation beyond the cytoplasmic membrane to be active. Reporters have their native export signal removed, and cannot be exported on their own. Instead, fused signal sequences or transmembrane domains are required to direct export of the reporter. When the protein-reporter fusion is exported, the activity of the enzyme is detected as proof of export. For example, the enzymatic activity may be "reported" on agar plates as CFU color change following cleavage of a colorimetric enzyme substrate. One reporter that has been used extensively to detect protein export in bacteria is alkaline phosphatase, or PhoA (20, 46, 57, 62, 63, 100). When PhoA has its native export signal removed ('PhoA) it cannot be exported and remains

intracellular. When fused to a heterologous export signal, the protein-'PhoA fusion is exported. When bacteria expressing such fusions are plated on agar containing the colorimetric PhoA substrate, the colonies turn blue. Thus, heterogeneous populations of bacteria can be screened for the presence of blue colonies in order to identify protein fusions capable of exporting of the 'PhoA reporter.

The 'PhoA reporter has been used to identify *M. tuberculosis* exported proteins. The first study to do so adapted the reporter for the identification of *M. tuberculosis* exported proteins using a genomic library approach (57). 'PhoA was encoded on a library vector downstream from a multiple cloning site. Genomic fragments from *M. tuberculosis* were cloned upstream of the reporter. The library was then electroporated into the non-pathogenic *M. smegmatis*, and resulting colonies were screened for blue color. Genetic fusions that allowed in-frame fusions of *M. tuberculosis* exported proteins to 'PhoA result in blue colonies. Out of 14,000 *M. smegmatis* library colonies, 12 express active and exported 'PhoA fusions. This study was conducted before the publication of the sequenced *M. tuberculosis* genome, so only three of the 12 sequences are identified as encoding exported proteins.

A second study incorporated '*phoA* onto a transposon, and used in vitro transposon mutagenesis to randomly insert the reporter in a *M. tuberculosis* cosmid library. The mutagenized cosmid library was then electroporated into *M. smegmatis* (10). Transposon insertion that allows 'PhoA expression in-frame with an exported protein results in blue colonies. In this study, 31 unique proteins capable of exporting active 'PhoA fusions are identified (10).

Using 'PhoA on transposons (TnPhoA) directly in pathogens has proven to be a powerful genetic approach for identifying exported virulence factors (46, 100). This is because the transposon simultaneously identifies exported proteins and knocks out the ORFS that encode them. However, the 'PhoA reporter cannot be used directly in virulent *M. tuberculosis* due to endogenous enzyme activities in the pathogen, which is why the approaches described above were carried out in *M. smegmatis*. Therefore, a reporter transposon system that identifies exported proteins directly in the pathogen is a desirable genetic tool. We describe such a tool in Chapter 4 of this dissertation.

Large-Scale Genetic Manipulation of Mycobacteria

In Chapters 2, 3 and 4 of this dissertation, I discuss genetic reporters and genome-wide strategies used to identify *M. tuberculosis* exported proteins. Our relatively limited understanding of the molecular mechanisms behind the ability of *M. tuberculosis* to cause disease is largely due to difficulties in studying this pathogen. These difficulties include a slow replication rate (24 hour doubling time), the requirement for work to be conducted in biosafety level three containment, and the slow development of genetic tools. In regards to this last point, over the last twelve years genetic tools for *M. tuberculosis* have increased in number and sophistication. The complete genome sequence of *M. tuberculosis* (21) is another significant resource now available. The result is that genomic-scale studies of *M. tuberculosis* are now feasible. *M. tuberculosis* has now been subjected to microarray analysis, signature tagged mutagenesis (STM), and saturating transposon mutagenesis all with the goal of identifying genes with intracellular or in vivo roles in promoting virulence (16, 22, 80, 90, 91). Experiments that globally analyze bacterial genes provide tremendous amounts of data in a relatively short period of time. However, results obtained from these experiments must be independently verified in more direct studies. The unifying theme of this dissertation research is the introduction of new tools for genome-wide identification of ORFS encoding exported proteins. Specifically, our goal was to identify new *M. tuberculosis* exported proteins with roles in virulence. Below is a brief discussion of three large-scale genomic analyses used to predict virulence factors in *M. tuberculosis*: Signature Tagged Mutagenesis (STM), Transposon Site Hybridization (TraSH) and Designer Arrays for Defined Mutant Analysis (DeADMAn). A discussion of the attributes and drawbacks associated with these types of analyses is also included.

Signature Tagged Mutagenesis (STM)

STM screens were the first to globally search for virulence factors in the *M. tuberculosis* genome. These mutagenesis screens were unquestionably successful because they immediately revealed important aspects of *M. tuberculosis* pathogenesis. In STM, a set of transposons is constructed in which each transposon has a unique oligonucleotide tag. *M. tuberculosis* is then subjected to transposon mutagenesis with each of the uniquely tagged transposons. Mice are then infected with mixed pools of tagged mutants. The tags can then be amplified from pools of insertion mutants before and after infection and detected with a radiolabeled set of probes. Transposon tags that are missing or underrepresented when compared to the input pool following hybridization to the probe indicate strains that are attenuated during infection. Attenuation is likely due to the loss of a potential virulence factor. These mutants are then recovered from the input pool and sequenced to reveal the site of transposon insertion. In this manner, the ORF predicted to be responsible for the observed phenotype is identified.

Two *M. tuberculosis* STM screens for virulence factors required during mouse infection were published simultaneously. The first examined an STM mutant library in the *M. tuberculosis* strain MT103 (16). A total of 1927 mutants were screened in pools of 48, delivered by intravenous injection. Out of these, 12 unique ORFS were identified as attenuated for virulence in the mouse lung. Interestingly, 9 of the 12 ORFs encode predicted exported proteins. The second STM screen (22) was conducted in the *M. tuberculosis* Erdman strain. Out of 576 mutants screened by tail vein injection, 14 are attenuated in mouse lung. Three of these mutants discussed in the published report overlap with attenuated strains identified in the other STM screen mentioned above (16). These overlapping mutants identified in both studies lack exported proteins and are defective in producing a surface lipid, phthiocerol dimycoserosate (PDIM) (38), required for virulence. The identity of the genes identified in these STM screens results underscore the importance of exported proteins in *M. tuberculosis* pathogenesis.

Transposon Site Hybridization (TraSH)

Another strategy used to determine the complement of genes required for growth in vitro, in macrophages or in mice, involves saturating transposon mutagenesis in *M. tuberculosis* and screening under different growth conditions using a technique called transposon site hybridization, or TraSH (80, 89-91). In these screens, the *M. tuberculosis* genome is subjected to transposon mutagenesis so that insertion in every gene is likely. A pooled library of transposon mutants is then allowed to grow under different conditions – in vitro, in macrophages, or in mouse spleens following intravenous injection. Sequences adjacent to the transposon insertion in each mutant are amplified from the pool of mutants before and after in vitro growth or infection. Following hybridization to an *M. tuberculosis* microarray, sequences present in the input pools can be compared with sequences represented in the output pools. Sequences missing or significantly depleted from the output pool represent genes that when disrupted attenuate the ability of the bacilli to replicate. This strategy serves to predict the genes that are required for optimal growth under the condition being studied.

During construction of the original TraSH mutant library in *M. tuberculosis*, it was observed that some ORFs did not receive insertions under saturating mutagenesis conditions, indicating that these ORFs encoded essential genes (90). It is predicted that a total of 614 genes, representing approximately 15% of the *M. tuberculosis* genome, are essential. Surprisingly, a third of these essential genes have no homology to inform on their function.

When the TraSH library of mutants was used to determine the suite of genes required for growth in mouse spleens, 194 *M. tuberculosis* genes were predicted to be required for in vivo survival and replication, representing approximately 5% of the genome (21, 91). Notably, 107 out of the 194 potential virulence factors, more than half, have no predicted function, and 25% of them are restricted to mycobacteria and closely related species. This data is indicative of how little is known about the mechanisms involved in *M. tuberculosis* pathogenesis. Once again, the majority of these proteins appear to be involved in protein export, exported protein synthesis, or are themselves exported proteins.

The TraSH strategy was also applied to a screen to identify mutants defective for growth in primary mouse macrophages (80). This screen was directed towards discovery of gene products required early in infection, as replication in macrophages is a necessary first step for spread of the pathogen in vivo. The macrophage study identified 126 genes as necessary for survival in macrophages. Both the TraSH in macrophages and TraSH in mice assays predict genes that encode proven *M. tuberculosis* virulence factors, indicating that these large-scale assays are useful for predicting genes that are important for pathogenesis.

Designer Arrays for Defined Mutant Analysis (DeADMAn)

The DeADMAn analysis takes advantage of a defined pool of transposon mutants. (51, 52). The insertion site of 1,425 *M. tuberculosis* CDC1551 transposon mutants were mapped, and those that fell upstream of the final 100 nucleotides of a gene were pooled and used in a screen to define genes required for growth in mouse lungs, livers and spleens following intravenous infection. This pre-selection made it more likely that the mutants included had transposon insertions that inactivated the respective ORF. Prior to infection, the pool of mutants was further selected based on comparable in vitro growth rates – all transposon mutants included in the in vivo screen were equally fit in growth medium. The mutant pools were then used to infect mice. Following infection, the tagged mutants were analyzed in a similar manner to the TraSH assays using microarray hybridization. Out of 530 unique transposon insertions in 514 unique genes, 31 transposon mutants in 30 unique genes were identified as defective for growth in mouse organs.

While large-scale genomic screens are powerful methods for providing biological information on a whole-organism level, there are drawbacks to these types of approaches. First, saturating mutagenesis screens for genes involved in pathogenesis will always miss those that are essential. Non-essential genes in operons that also contain essential genes may also be missed due to polar effects. While

essential genes are often required for metabolic function and may not represent new classes of virulence factors, these families should not be discounted, especially in a pathogen whose only known niche is in the cells of infected humans and animals. Further, essential genes are excellent targets for novel drug therapies against tuberculosis.

Second, successful saturation transposon mutagenesis studies rely on the assumption that every ORF is equally mutable and that in the transposition reactions conducted in these studies transposition has occurred in every mutable ORF. While transposons used in *M. tuberculosis* can insert in a variety of sites in the genome, no transposon is completely random, and again, essential genes, as well as operons with essential genes, will be missed. Further, there are *M. tuberculosis* ORFs that are not present on the microarrays used in the DeADMAn and TraSH screens. No data are can be obtained about ORFs that are absent from the microarray during these types of large-scale infection screens.

Another draw back to genome-wide screens is that they must rely on arbitrary cut-offs when comparing ratios of representation in input pools vs. output pools. Cut-offs are necessary to organize the massive amount of data that results from high-throughput screens. All genes that are underrepresented below a certain cut-off ratio in the output pools are considered defective, but there will always be some degree of false negative and false positive findings. Careful follow up experiments are required to validate results obtained from screening large pools of mutants. Due to the potential for polar effects when using transposon mutagenesis, complementation experiments are also required to demonstrate that the mutated ORF is indeed the cause of the observed phenotype.

Infection with pools allows a large number of mutants to be screened simultaneously for phenotypes in small number of animals. However, there are known bottlenecks that occur during mouse infection. Following delivery of *M. tuberculosis* to mice by the intravenous route, approximately 90% of bacteria are found in the liver, 9% are in the spleen, and 1% is in the lungs (71). Therefore, it is possible that mutants may appear to be underrepresented in high throughput screens when in fact they were never seeded to the lungs. Also, methods that evaluate pools of mutants may miss important genes

due to masking by complementation by other bacteria in trans. This is a particular concern for genes encoding exported proteins. A co-infecting bacterium that exports a virulence factor may rescue a mutant lacking that virulence factor. In this manner, attenuated mutants might be overlooked.

Finally, no single genome-wide screen will be able to identify all virulence factors. The results obtained from genome wide screens are dictated by the study design and the types of scientific questions asked (31). Depending on route of infection, organ infected, model system and the chosen time point post infection, different sets of genes will be predicted to be important. Further, there are mutants that exhibit no growth defect in organs, but do elicit different tissue pathology and different outcomes in long-term survival (47, 99). Such mutants would be missed in the STM, TraSH and DeADMAn assays. However, high-throughput screens are useful in that they provide snapshots of *M. tuberculosis* biology as a whole as it operates under certain conditions. Data resulting from large-scale genomic screens should be viewed in this context.

Summary

As in other bacterial pathogens, exported proteins of *M. tuberculosis* are sure to play important roles in virulence. The relatively recent ability to construct *M. tuberculosis* mutants has greatly advanced our understanding of the protein transport systems and the importance of exported proteins to virulence of this pathogen. However, at this point, our understanding of the function of even the best studied virulence factors of *M. tuberculosis* remains minimal. Future investigations are sure to focus more directly on how proteins exported by this pathogen interact with the host and cause disease.

The focus of this dissertation is the development of new methods to identify and study exported proteins of *M. tuberculosis*. For many of the proteins identified in this research, these studies provide the first experimental evidence of export. Chapter 2 will describe the use of a new reporter system to specifically identify *M. tuberculosis* proteins exported by the Tat pathway. Chapter 3 will describe a new reporter system with the unique ability to report on protein export during *M. tuberculosis* growth in

host cells. Chapter 4 will describe a new reporter transposon used to simultaneously identify exported proteins of *M. tuberculosis* and collect mutants lacking these proteins. The resulting library of *M. tuberculosis* mutants was successfully screened for mutants defective for growth in macrophages as a strategy to identify new exported virulence factors.

REFERENCES

1. **Andersen, P., A. B. Andersen, A. L. Sorensen, and S. Nagai.** 1995. Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice. J Immunol **154:**3359-72.

2. **Archambaud, C., M. A. Nahori, J. Pizarro-Cerda, P. Cossart, and O. Dussurget.** 2006. Control of *Listeria* superoxide dismutase by phosphorylation. J Biol Chem.

3. **Behr, M. A., and P. M. Small.** 1999. A historical and molecular phylogeny of BCG strains. Vaccine **17:**915-22.

4. **Bensing, B. A., and P. M. Sullam.** 2002. An accessory *sec* locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. Mol Microbiol **44:**1081-94.

5. **Berthet, F. X., M. Lagranderie, P. Gounon, C. LaurentWinter, D. Ensergueix, P. Chavarot, F. Thouron, E. Maranghi, V. Pelicic, D. Portnoi, G. Marchal, and B. Gicquel.** 1998. Attenuation of virulence by disruption of the *Mycobacterium tuberculosis erp* gene. Science **282:**759-762.

6. **Berthet, F. X., J. Rauzier, E. M. Lim, W. Philipp, B. Gicquel, and D. Portnoi.** 1995. Characterization of the *Mycobacterium tuberculosis erp* gene encoding a potential cell surface protein with repetitive structures. Microbiology **141:**2123-30.

7. **Braunstein, M., and J. T. Belisle.** 2000. Genetics of protein secretion, p. 203-220. *In* G. F. Hatfull and W. R. J. Jacobs (ed.), Molecular genetics of mycobacteria. ASM Press.

8. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-90.

9. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Molecular Microbiology **48:**453-464.

10. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552'*phoA* in vitro transposition system. J Bacteriol **182:**2732-40.

11. **Brennan, P. J., and H. Nikaido.** 1995. The envelope of mycobacteria. Annu Rev Biochem **64:**29-63.

12. **Brodin, P., M. I. de Jonge, L. Majlessi, C. Leclerc, M. Nilges, S. T. Cole, and R. Brosch.** 2005. Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis,* reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. J Biol Chem **280:**33953-9.

13. **Brodin, P., L. Majlessi, L. Marsollier, M. I. de Jonge, D. Bottai, C. Demangel, J. Hinds, O. Neyrolles, P. D. Butcher, C. Leclerc, S. T. Cole, and R. Brosch.** 2006. Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. Infect Immun **74:**88-98.

14. **Burts, M. L., W. A. Williams, K. DeBord, and D. M. Missiakas.** 2005. EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. Proc Natl Acad Sci U S A **102:**1169-74.

15. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. Mol Plant Microbe Interact **19:**200-12.

16. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of mycobacterium tuberculosis by signature-tagged transposon mutagenesis. Mol Microbiol **34:**257-67.

17. **Champion, P. A., S. A. Stanley, M. M. Champion, E. J. Brown, and J. S. Cox.** 2006. C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. Science **313:**1632-6.

18. **Chen, Q., H. Wu, and P. M. Fives-Taylor.** 2004. Investigating the role of *secA2* in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. Mol Microbiol **53:**843-56.

19. **Christie, P. J.** 2001. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol Microbiol **40:**294-305.

20. **Cleavinger, C. M., M. F. Kim, J. H. Im, and K. S. Wise.** 1995. Identification of mycoplasma membrane proteins by systematic Tn *phoA* mutagenesis of a recombinant library. Mol Microbiol **18:**283-93.

21. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-44.

22. **Cox, J. S., B. Chen, M. McNeil, and W. R. Jacobs, Jr.** 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature **402:**79-83.

23. **Daffe, M., and P. Draper.** 1998. The envelope layers of mycobacteria with reference to their pathogenicity. Adv Microb Physiol **39:**131-203.

24. **Dave, J. A., N. C. Gey van Pittius, A. D. Beyers, M. R. Ehlers, and G. D. Brown.** 2002. Mycosin-1, a subtilisin-like serine protease of *Mycobacterium tuberculosis*, is cell wall-associated and expressed during infection of macrophages. BMC Microbiol **2:**30.

25. **De Buck, E., L. Maes, E. Meyen, L. Van Mellaert, N. Geukens, J. Anne, and E. Lammertyn.** 2005. *Legionella pneumophila* Philadelphia-1 *tatB* and *tatC* affect intracellular replication and biofilm formation. Biochem Biophys Res Commun **331:**1413-20.

26. **Demangel, C., P. Brodin, P. J. Cockle, R. Brosch, L. Majlessi, C. Leclerc, and S. T. Cole.** 2004. Cell envelope protein PPE68 contributes to *Mycobacterium tuberculosis* RD1

immunogenicity independently of a 10-kilodalton culture filtrate protein and ESAT-6. Infect Immun **72:**2170-6.

27. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-83.

28. **Dye, C., S. Scheele, P. Dolin, V. Pathania, and M. C. Raviglione.** 1999. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. Jama **282:**677-86.

29. **Emanuelsson, O., S. Brunak, G. von Heijne, and H. Nielsen.** 2007. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc **2:**953-71.

30. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev **61:**136-69.

31. **Fortune, S. M., M. R. Chase, and E. J. Rubin.** 2006. Dividing oceans into pools: strategies for the global analysis of bacterial genes. Microbes Infect **8:**1631-6.

32. **Fortune, S. M., A. Jaeger, D. A. Sarracino, M. R. Chase, C. M. Sassetti, D. R. Sherman, B. R. Bloom, and E. J. Rubin.** 2005. Mutually dependent secretion of proteins required for mycobacterial virulence. Proc Natl Acad Sci U S A **102:**10676-81.

33. **Fratazzi, C., R. D. Arbeit, C. Carini, M. K. Balcewicz-Sablinska, J. Keane, H. Kornfeld, and H. G. Remold.** 1999. Macrophage apoptosis in mycobacterial infections. J Leukoc Biol **66:**763-4.

34. **Frieden, T. R., and G. R. Khatri.** 2003. Impact of national consultants on successful expansion of effective tuberculosis control in India. Int J Tuberc Lung Dis **7:**837-41.

35. **Gey Van Pittius, N. C., J. Gamieldien, W. Hide, G. D. Brown, R. J. Siezen, and A. D. Beyers.** 2001. The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. Genome Biol **2:**RESEARCH0044.

36. **Gey van Pittius, N. C., S. L. Sampson, H. Lee, Y. Kim, P. D. van Helden, and R. M. Warren.** 2006. Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. BMC Evol Biol **6:**95.

37. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol.

38. **Guilhot, C. a. D., M.** 2008. Polyketides and Polyketide-Containing Glycolipids of *Mycobacterium tuberculosis*: Structure, Biosynthesis and Biological Activities. *In* S. H. a. R. Kaufmann, E. (ed.), Handbook of Tuberculosis. Wiley VCH, Weinheim, Germany.

39. **Guinn, K. M., M. J. Hickey, S. K. Mathur, K. L. Zakel, J. E. Grotzke, D. M. Lewinsohn, S. Smith, and D. R. Sherman.** 2004. Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. Mol Microbiol **51:**359-70.

40. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing essential protein secretion in *Mycobacterium smegmatis* using tetracycline repressors. J Bacteriol.

41. **Harboe, M., T. Oettinger, H. G. Wiker, I. Rosenkrands, and P. Andersen.** 1996. Evidence for occurrence of the ESAT-6 protein in *Mycobacterium tuberculosis* and virulent *Mycobacterium bovis* and for its absence in *Mycobacterium bovis* BCG. Infect Immun **64:**16-22.

42. **Hou, J. M., N. G. D'Lima, N. W. Rigel, H. S. Gibbons, J. R. McCann, M. Braunstein, and C. M. Teschke.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-7.

43. **Houben, E. N., L. Nguyen, and J. Pieters.** 2006. Interaction of pathogenic mycobacteria with the host immune system. Curr Opin Microbiol **9:**76-85.

44. **Hsu, T., S. M. Hingley-Wilson, B. Chen, M. Chen, A. Z. Dai, P. M. Morin, C. B. Marks, J. Padiyar, C. Goulding, M. Gingery, D. Eisenberg, R. G. Russell, S. C. Derrick, F. M. Collins, S. L. Morris, C. H. King, and W. R. Jacobs, Jr.** 2003. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. Proc Natl Acad Sci U S A **100:**12420-5.

45. **Johansen, K. A., R. E. Gill, and M. L. Vasil.** 1996. Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. Infect Immun **64:**3259-66.

46. **Kaufman, M. R., and R. K. Taylor.** 1994. Identification of bacterial cell-surface virulence determinants with Tn*phoA*. *In* V. L. Clark and P. M. Bavoil (ed.), Methods in Enzymology, vol. 235. Academic Press, Inc., New York.

47. **Kaushal, D., B. G. Schroeder, S. Tyagi, T. Yoshimatsu, C. Scott, C. Ko, L. Carpenter, J. Mehrotra, Y. C. Manabe, R. D. Fleischmann, and W. R. Bishai.** 2002. Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH. Proc Natl Acad Sci U S A **99:**8330-5.

48. **Kurtz, S., and M. Braunstein.** 2005. Protein secretion and export in *Mycobacterium tuberculosis*, p. 71-138. *In* T. Parish (ed.), *Mycobacterium* molecular biology. Horizon bioscience, Norfolk, UK.

49. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun.

50. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun **74:**6855-64.

51. **Lamichhane, G., S. Tyagi, and W. R. Bishai.** 2005. Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. Infect Immun **73:**2533-40.

52. **Lamichhane, G., M. Zignol, N. J. Blades, D. E. Geiman, A. Dougherty, J. Grosset, K. W. Broman, and W. R. Bishai.** 2003. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **100:**7213-8.

53. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. Infect Immun **74:**1768-76.

54. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. Annu Rev Microbiol **60:**373-395.

55. **Lenz, L. L., and D. A. Portnoy.** 2002. Identification of a second *Listeria* secA gene associated with protein secretion and the rough phenotype. Mol Microbiol **45:**1043-56.

56. **Lewis, K. N., R. Liao, K. M. Guinn, M. J. Hickey, S. Smith, M. A. Behr, and D. R. Sherman.** 2003. Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. J Infect Dis **187:**117-23.

57. **Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi.** 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. J Bacteriol **177:**59-65.

58. **Luthra, A., A. Mahmood, A. Arora, and R. Ramachandran.** 2008. Characterization of Rv3868, an essential hypothetical protein of the ESX-1 secretion system in *Mycobacterium tuberculosis*. J Biol Chem **283:**36532-41.

59. **MacGurn, J. A., S. Raghavan, S. A. Stanley, and J. S. Cox.** 2005. A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*. Mol Microbiol **57:**1653-63.

60. **Mahairas, G. G., P. J. Sabo, M. J. Hickey, D. C. Singh, and C. K. Stover.** 1996. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. J Bacteriol **178:**1274-82.

61. **Malen, H., F. S. Berven, K. E. Fladmark, and H. G. Wiker.** 2007. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosi*s H37Rv. Proteomics **7:**1702-18.

62. **Manoil, C., and J. Beckwith.** 1985. TnphoA: a transposon probe for protein export signals. Proc Natl Acad Sci U S A **82:**8129-33.

63. **Manoil, C., J. J. Mekalanos, and J. Beckwith.** 1990. Alkaline phosphatase fusions: sensors of subcellular location. J Bacteriol **172:**515-8.

64. **Mattow, J., U. E. Schaible, F. Schmidt, K. Hagens, F. Siejak, G. Brestrich, G. Haeselbarth, E. C. Muller, P. R. Jungblut, and S. H. Kaufmann.** 2003. Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated M. bovis BCG Copenhagen. Electrophoresis **24:**3405-20.

65. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-79.

66. **McLaughlin, B., J. S. Chon, J. A. MacGurn, F. Carlsson, T. L. Cheng, J. S. Cox, and E. J. Brown.** 2007. A mycobacterium ESX-1-secreted virulence factor with unique requirements for export. PLoS Pathog **3:**e105.

67. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. Trends Microbiol **9:**494-500.

68. **Ochsner, U. A., A. Snyder, A. I. Vasil, and M. L. Vasil.** 2002. Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. Proc Natl Acad Sci U S A **99:**8312-7.

69. **Okkels, L. M., and P. Andersen.** 2004. Protein-protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis*. J Bacteriol **186:**2487-91.

70. **Okkels, L. M., I. Brock, F. Follmann, E. M. Agger, S. M. Arend, T. H. Ottenhoff, F. Oftung, I. Rosenkrands, and P. Andersen.** 2003. PPE protein (Rv3873) from DNA segment RD1 of *Mycobacterium tuberculosis*: strong recognition of both specific T-cell epitopes and epitopes conserved within the PPE family. Infect Immun **71:**6116-23.

71. **Orme, I., and F. M. Collins.** 1994. Mouse model of tuberculosis. *In* B. Bloom (ed.), Tuberculosis. American Society for Microbiology, Washington D.C.

72. **Pallen, M. J.** 2002. The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? Trends Microbiol **10:**209-12.

73. **Pan, H., B. S. Yan, M. Rojas, Y. V. Shebzukhov, H. Zhou, L. Kobzik, D. E. Higgins, M. J. Daly, B. R. Bloom, and I. Kramnik.** 2005. Ipr1 gene mediates innate immunity to tuberculosis. Nature **434:**767-72.

74. **Papanikou, E., S. Karamanou, and A. Economou.** 2007. Bacterial protein secretion through the translocase nanomachine. Nat Rev Microbiol **5:**839-51.

75. **Plano, G. V., J. B. Day, and F. Ferracci.** 2001. Type III export: new uses for an old pathway. Mol Microbiol **40:**284-93.

76. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-40.

77. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71:**4908-4916.

78. **Pym, A. S., P. Brodin, L. Majlessi, R. Brosch, C. Demangel, A. Williams, K. E. Griffiths, G. Marchal, C. Leclerc, and S. T. Cole.** 2003. Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. Nat Med **9:**533-9.

79. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-17.

80. **Rengarajan, J., B. R. Bloom, and E. J. Rubin.** 2005. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A **102:**8327-32.

81. **Renshaw, P. S., K. L. Lightbody, V. Veverka, F. W. Muskett, G. Kelly, T. A. Frenkiel, S. V. Gordon, R. G. Hewinson, B. Burke, J. Norman, R. A. Williamson, and M. D. Carr.** 2005. Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. Embo J **24:**2491-8.

82. **Renshaw, P. S., P. Panagiotidou, A. Whelan, S. V. Gordon, R. G. Hewinson, R. A. Williamson, and M. D. Carr.** 2002. Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6*CFP-10 complex. Implications for pathogenesis and virulence. J Biol Chem **277:**21598-603.

83. **Rhee, K. Y.** 2005. Mycobacterial resistance to reactive oxygen and nitrogen intermediates: recent views and progress in *M. tuberculosis*, p. 321-346. *In* T. Parish (ed.), *Mycobacterium* molecular methods. Horizon bioscience.

84. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The Accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-36.

85. **Rosenkrands, I., A. King, K. Weldingh, M. Moniatte, E. Moertz, and P. Andersen.** 2000. Towards the proteome of *Mycobacterium tuberculosis*. Electrophoresis **21:**3740-56.

86. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? Nat Rev Microbiol **5:**39-47.

87. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-79.

88. **Sandkvist, M.** 2001. Type II secretion and pathogenesis. Infect Immun **69:**3523-35.

89. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc Natl Acad Sci U S A **98:**12712-7.

90. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

91. **Sassetti, C. M., and E. J. Rubin.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-94.

92. **Schneider, G.** 1999. How many potentially secreted proteins are contained in a bacterial genome? Gene **237:**1113-1121.

93. **Singh, A., D. Mai, A. Kumar, and A. J. Steyn.** 2006. Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. Proc Natl Acad Sci U S A **103:**11346-51.

94. **Skjot, R. L., I. Brock, S. M. Arend, M. E. Munk, M. Theisen, T. H. Ottenhoff, and P. Andersen.** 2002. Epitope mapping of the immunodominant antigen TB10.4 and the two homologous proteins TB10.3 and TB12.9, which constitute a subfamily of the *esat-6* gene family. Infect Immun **70:**5446-53.

95. **Snyder, A., A. I. Vasil, S. L. Zajdowicz, Z. R. Wilson, and M. L. Vasil.** 2006. Role of the *Pseudomonas aeruginosa* PlcH Tat signal peptide in protein secretion, transcription, and cross-species Tat secretion system compatibility. J Bacteriol **188:**1762-74.

96. **Sorensen, A. L., S. Nagai, G. Houen, P. Andersen, and A. B. Andersen.** 1995. Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. Infect Immun **63:**1710-7.

97. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli. J Biol Chem **275:**11591-6.

98. **Stanley, S. A., S. Raghavan, W. W. Hwang, and J. S. Cox.** 2003. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. Proc Natl Acad Sci U S A **100:**13001-6.

99. **Steyn, A. J., D. M. Collins, M. K. Hondalus, W. R. Jacobs, Jr., R. P. Kawakami, and B. R. Bloom.** 2002. *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. Proc Natl Acad Sci U S A **99:**3147-3152.

100. **Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos.** 1987. Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc Natl Acad Sci U S A **84:**2833-7.

101. **Tekaia, F., S. V. Gordon, T. Garnier, R. Brosch, B. G. Barrell, and S. T. Cole.** 1999. Analysis of the proteome of *Mycobacterium tuberculosis in silico*. Tuber Lung Dis **79:**329-42.

102. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. J Biol Chem.

103. **Vergne, I., J. Chua, S. B. Singh, and V. Deretic.** 2004. Cell biology of *mycobacterium tuberculosis* phagosome. Annu Rev Cell Dev Biol **20:**367-94.

104. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. Embo J **20:**6735-41.

105. **WorldHealthOrganization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.] http://www.who.int/mediacentre/factsheets/fs104/en/

106. **WorldHealthOrganization.** 2007. XDR-TB; What, where, how and action steps. http://www.who.int/mediacentre/news/notes/2006/np23/en/index.html

CHAPTER 2

Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins

Justin A. McDonough, Jessica R. McCann, Erin McElvania Tekippe,
Jason S. Silverman, Nathan W. Rigel, and Miriam Braunstein

The twin-arginine translocation (Tat) pathway is a system used by some bacteria to export proteins out from the cytosol to the cell surface or extracellular environment. A functional Tat pathway exists in the important human pathogen, *Mycobacterium tuberculosis*. Identification of the substrates exported by the Tat pathway can help define the role this pathway plays in the physiology and pathogenesis of *M. tuberculosis*. Here we used a reporter of Tat export, a truncated β-lactamase 'BlaC, to experimentally identify *M. tuberculosis* proteins with functional Tat signal sequences. Of the 13 proteins identified one lacks the hallmark of a Tat exported substrate, the twin-arginine dipeptide, and another is not predicted by *in silico* analysis of the annotated *M. tuberculosis* genome. Full-length versions of a subset of these proteins were tested to determine if the native proteins are Tat exported. For three proteins, expression in a Δ*tat* mutant of *M. smegmatis* revealed a defect in precursor processing compared to expression in wild-type indicating Tat-export of the full-length proteins. Conversely, two proteins showed no obvious Tat export in *M. smegmatis*. One of this latter group of proteins was the *M. tuberculosis* virulence factor phospholipase C, PlcB. Importantly, when tested in *M. tuberculosis* a different result was obtained and PlcB was exported in a twin-arginine dependent manner. This suggests the existence of *M. tuberculosis*-specific factor(s) for Tat export of a proven virulence

protein. It also emphasizes the importance of domains beyond the Tat signal sequence and bacteria-specific factors in determining if a given protein is Tat-exported.

INTRODUCTION

In bacteria, protein export across the cytoplasmic membrane represents the first step in the delivery of proteins to the cell envelope or extracellular space. Two conserved systems are responsible for the majority of this protein export: the general secretion (Sec) and the twin-arginine translocation (Tat) pathways (for reviews, see (31, 35). Both systems export proteins that are synthesized as precursors with amino-terminal signal sequences. In both cases, the signal sequences are comprised of a tripartite structure: a charged amino-terminal region, a hydrophobic region, and a carboxy-terminal region containing a signal peptidase cleavage site (31, 62). With most exported proteins, the signal sequence is cleaved from the precursor during or immediately after translocation which liberates the mature exported substrate.

A feature that distinguishes Tat signal sequences from Sec signal sequences is the presence of a consensus twin-arginine motif, which is defined as S/T-R-R-x-F-L-K (5). The arginine dipeptide 'RR' in the motif is a major targeting determinant of the signal sequence as shown by conservative substitution of 'RR' with a lysine pair 'KK' preventing Tat-export of proteins (13, 14, 25, 54). Computational Tat signal sequence prediction programs, based on sequence and structural conservation, have been developed (4, 45, 52). These programs are valuable for identifying Tat substrates that adhere to the consensus motif; however, they cannot account for species specific differences, unless modified to do so, and it remains to be established how useful they are for comprehensive identification of Tat exported proteins. There are two Tat-exported proteins known to exist in nature that lack the twin-arginines (26, 27). These exceptions may be members of a larger group of yet-to-be identified Tat proteins that rely on features other than the twin-arginine motif for Tat export.

Another distinguishing feature of the Tat export pathway is that Tat substrates are translocated across the membrane in a folded state, with folding being a prerequisite for Tat export (15). Some Tat

substrates require cytoplasmic chaperones for export. These chaperones may be specific to one Tat substrate or they can have a more general effect (23, 28, 38-40). In these cases, it is thought that chaperones function in folding substrates or targeting them to the membrane localized Tat translocase complex once folding is complete. The Tat translocase is composed of TatA, Tat B and TatC proteins, although not all bacteria with functional Tat export systems have TatB.

The Tat pathway is present in many, but not all, bacteria. In several bacterial pathogens, the Tat pathway plays an important role in exporting virulence factors (9, 10, 17, 30, 36, 42, 46, 60). *Mycobacterium tuberculosis* is the bacterial pathogen responsible for tuberculosis, which kills 1.8 million people a year (64). Mycobacteria have a functional Tat pathway. In the fast-growing non-pathogenic *Mycobacterium smegmatis*, mutants lacking *tatA*, *tatB*, or *tatC* genes have multiple phenotypes including slow growth on agar and sensitivity to β-lactam antibiotics (34, 41). The latter phenotype is attributed to a failure to export the chromosomally-encoded β-lactamase. In both *M. smegmatis* and *M. tuberculosis* the endogenous β-lactamases possess Tat signal sequences. β-lactamases, which destroy cell wall-targeting β-lactam antibiotics, must be exported to protect bacteria from the drugs. In *M. tuberculosis* it has not been possible to construct *tat* mutants (47). This indicates that, in pathogenic *M. tuberculosis*, the Tat pathway is essential under standard laboratory conditions. Without a *M. tuberculosis tat* mutant, there are fewer approaches available for identifying Tat exported proteins and studying the significance of Tat export in this pathogen.

In this study, we used the *M. tuberculosis* β-lactamase (BlaC) as a reporter to identify *M. tuberculosis* proteins that possess functional Tat signal sequences. A truncated ‘BlaC, lacking its endogenous signal sequence, is not exported and is unable to protect a mycobacterial β-lactam-sensitive mutant (*M. smegmatis* Δ*blaS* or *M. tuberculosis* Δ*blaC*) from the β-lactam antibiotic carbenicillin (34). When a signal sequence from a Tat exported *M. tuberculosis* protein is fused to ‘BlaC, the hybrid protein is exported and confers carbenicillin resistance to Δ*bla* mutant mycobacteria. Exported ‘BlaC fusion proteins can be identified by direct selection of drug resistant colonies on agar containing

carbenicillin. Importantly, the 'BlaC reporter is Tat specific. It works only when fused to Tat signal sequences and requires both the twin-arginine motif and a functional Tat pathway (34).

Using a *M. tuberculosis* genomic library constructed upstream of the *'blaC* reporter, we identified signal sequences capable of exporting 'BlaC in a Tat-dependent manner. In addition to the demonstrated virulence factor and phospholipase C (PlcB) (29, 43), we identified proteins with potential roles in carbohydrate and lipid metabolism, copper homeostasis, cell envelope maintenance, and nutrient import. The proteins identified included one lacking a twin-arginine dipeptide and one not predicted by *in silico* analysis. We also investigated full-length versions of a subset of the proteins identified. Importantly, full-length PlcB was exported and twin-arginine dependent when expressed in its native host *M. tuberculosis*. However, when expressed in *M. smegmatis* this *M. tuberculosis* protein did not appear to be exported. This suggests the existence of *M. tuberculosis*-specific factor(s) that are required for Tat-export of a proven virulence protein.

MATERIALS AND METHODS

Bacterial strains and culture methods. Bacterial strains used during this work are listed in Table 2.1. Luria-Bertani (LB) medium (Fisher) was used for culturing of *Escherichia coli*. Middlebrook 7H9 or 7H10 medium (Difco; BD Biosciences) was used for the culturing of *M. smegmatis* and *M. tuberculosis*. For *M. smegmatis*, Middlebrook medium was supplemented with 0.5% glycerol and 0.2% dextrose. For *M. tuberculosis*, Middlebrook medium was supplemented with 0.5% glycerol and 1× ADS (0.5% bovine serum albumin, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl). When necessary, media was supplemented with 0.05-0.1% Tween-80 (Fisher). As required, antibiotics were added to Middlebrook media at the following concentrations: hygromycin B (Roche Applied Science) 50 μg/ml; carbenicillin (Sigma), 50 μg/ml; or kanamycin (Acros Chemicals), 20 μg/ml. L-lysine at 40 and 80 μg/ml was added to agar and liquid media, respectively, for growth of *M. smegmatis* strains PM759 and JM578.

Molecular biology procedures. Standard molecular biology techniques were employed (48). The Expand high fidelity PCR system (Roche) was used in all PCR reactions and 5.0% dimethyl sulfoxide

Table 2.1. Strains used in this study

Strain	Description	Reference
E. coli		
K-12 DH5α	F- [ϕ80dΔ*lacZ*M15] Δ(*lacZYA-argF*) *U169 deoR recA1 endA1 hsdR17 glnV44 thi-1 gyrA96 relA1*	Gibco-BRL
TOP10	F- mcrA [ϕ80dΔ*lacZ*M15] Δ(*mrr-hsdRMS-mcrBC*) Δ*lacX74 deoR recA1 endA1 araD139* Δ(*ara, leu*)7697 *galU galK rpsL nupG*	Invitrogen
XL-1 Blue	[F' *proAB*, *lacI*^{q}*Z*Δ *M15*, Tn10(Tet^r)] *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1* λ$^-$ *lac*$^-$	Stratagene
M. smegmatis		
mc^{2}155	*ept-1*	(53)
MB692	mc^{2}155, Δ*tatA*	(34)
JM567	mc^{2}155, Δ*tatC*	(34)
PM759	mc^{2}155, Δ*blaS1* Δ*lysA4 rpsL6*	(18)
JM578	PM759, Δ*tatA*	(34)
M. tuberculosis		
H37Rv	Virulent	(1)
PM638	H37Rv, Δ*blaC1*	(18)

was included in select reactions. DNA sequencing was performed by either the UNC-CH Automated DNA Sequencing Facility (Chapel Hill, NC) or by Eton Bioscience Inc. (San Diego, CA).

Construction of 'BlaC reporter libraries. All plasmids used in this study are listed in Table 2.2. Two library plasmids were constructed with a truncated version of M. tuberculosis blaC (referred to as 'blaC), amplified from pJM106 by PCR using the primers Lib'blaCfor and Lib'blaCrev (Supplemental Table 2.1).

(i) Library 1. The resulting '*blaC* amplicon was ligated into the multicopy mycobacterial shuttle vector pMV206.*hyg*, which had been digested with *Cla*I and *Nco*I. The final plasmid was pJES113 (Fig. 2.1A). Genomic DNA was isolated from *M. tuberculosis* strain H37Rv as previously described (6), partially digested with *Aci*I and *Hpa*II, and digests with DNA fragments between 0.5 and 5.0-kbp were selected.

Table 2.2. Plasmids used in this study.

Plasmid	Genotype	Description	Source
pCR2.1	*bla aph* ColE1	TA cloning vector	Invitrogen
pMV206.*hyg*	*hyg oriM* ColE1	Multicopy mycobacterial shuttle plasmid	(55)
pMV261.*kan*	*aph* P_{hsp60} *oriM* ColE1	Multicopy mycobacterial shuttle plasmid	(55)
pMV306.*kan*	*aph int attP* ColE1	Single-copy mycobacterial shuttle plasmid	(55)
pJSC77	*aph* P_{hsp60} *-HA oriM* ColE1	HA tag cloned into pMV261	(21)
pJM106	*cat 'blaC (M. tuberculosis) oriV ori2*	Predicted *M. tuberculosis 'blaC* mature sequence cloned into pCC1 (Epicentre)	(34)
pJES112	*bla aph* ColE1 *'blaC* (*M. tuberculosis*)	*M. tuberculosis 'blaC* mature sequence cloned into pCR2.1	This work
pJES113	*hyg oriM* ColE1 *'blaC* (*M. tuberculosis*)	*M. tuberculosis 'blaC* mature sequence cloned into pMV206.hyg	This work
pJM153	*bla aph* ColE1P_{hsp60} (*M. tuberculosis*)	*M. tuberculosis hsp60* promoter cloned into pCR2.1	This work
pJM157	*hyg oriM* ColE1 P_{hsp60}, *'blaC* (*M. tuberculosis*)	*M. tuberculosis hsp60* promoter cloned into pJES113	This work
Y49 pTrcHisB-BlaC	*bla* P_{trc}*-blaC* (*M. tuberculosis*) ColE1	*E. coli* expression vector for *M. tuberculosis blaC*	(59)
pJSS51	*bla aph* ColE1 '*plcB* (*M. tuberculosis*)	*M. tuberculosis plcB* (no signal sequence) cloned into pCR2.1	This work
pJM171	*aph oriM* ColE1 P_{hsp60} '*plcB-HA*	1.4-kbp *Msc*I/*Hin*dIII fragment containing *plcB* (no signal sequence) from pJSS51 cloned into pJSC77	This work
pJM210	*bla aph* ColE1 *Rv0774c* (*M. tuberculosis*)	*M. tuberculosis* full length *Rv0774c* cloned into pCR2.1	This work
pJM211	*aph oriM* ColE1 P_{hsp60} *Rv0774c-HA*	913-bp *Msc*I/*Hin*dIII fragment containing full length *Rv0774c* from pJM210 cloned into pJSC77	This work
pJES146	*bla aph* ColE1	*M. tuberculosis Rv0774c* (no signal sequence) cloned into pCR2.1	This work
PJES147	*aph* P_{hsp60}'*Rv0774c -HA oriM* ColE1	797-bp *Msc*I/*Hin*dIII fragment containing Rv0774c (no signal sequence) from pJES146 cloned into pJSC77	This work
pJM212	*bla aph* ColE1 *Rv2843* (*M. tuberculosis*)	*M. tuberculosis* full length *Rv2843* cloned into pCR2.1	This work
pJM214	*aph oriM* ColE1 P_{hsp60} *Rv2843-HA*	547-bp *Nru*I/*Hin*dIII fragment containing full length *Rv2843* from pJM212 cloned into pJSC77	This work
pJM196	*bla aph* ColE1 P_{native} *plcB* (*M. tuberculosis*)	*M. tuberculosis* full length *plcB* and promoter cloned into pCR2.1	This work
pJM199	*aph oriM* ColE1 P_{native} *plcB-HA*	2.2-kbp *Xba*I/*Hin*dIII fragment containing full length *plcB* and promoter from pJM196 cloned into pJSC77	This work
pJM202	*bla aph* ColE1 P_{native} *plcA*, *plcB* (*M. tuberculosis*)	*M. tuberculosis* full length *plcA*, *plcB*, and promoter cloned into pCR2.1	This work
pJM203	*aph oriM* ColE1 P_{native} *plcA*, *plcB-HA*	3.5-kbp *Xba*I/*Hin*dIII fragment containing full length *plcA*, *plcB*, and promoter from pJM202 cloned into pJSC77	This work
pJM197	*bla aph* ColE1 P_{native} *Rv0315* (*M. tuberculosis*)	*M. tuberculosis* full length *Rv0315* and promoter cloned into pCR2.1	This work
pJM206	*aph oriM* ColE1 P_{native} *Rv0315-HA*	1.2-kbp *Xba*I/*Hin*dIII fragment containing full length *Rv0315* and promoter from pJM197 cloned into pJSC77	This work
pJM216	*aph oriM* ColEI P_{plcB}, *plcB(KK)-HA*	Overlapping, self-ligated PCR product amplified from pJM199. Contains *plcB-KK* under control of the native promoter	This work
pMP327	*aph* P_{blaC}*-blaC* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis* full length *blaC* cloned into pMV261	(34)
pJM113	*aph* P_{hsp60}*-'blaC* (*M. tuberculosis*) *oriM* ColE1	'*blaC* from pJM106 cloned into *Msc*I-linearized pMV261	(34)
pJM117	*aph* P_{blaC}*-blaC*(KK) (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis* full length *blaC*(KK) cloned into pMV261	(34)

The genomic digest was cloned into the unique *Cla*I site immediately upstream of '*blaC* in pJES113. The resulting ligation reaction was transformed into *E. coli* XL-1 Blue (Stratagene). Approximately 1×10^6 hygromycin-resistant *E. coli* transformants were pooled for plasmid DNA isolation (Qiagen). **(ii) Library 2.** The second library vector, pJM157, was constructed to carry a mycobacterial promoter upstream of the unique *Cla*I site. A fragment carrying the promoter and the +1 transcriptional start site from *M. tuberculosis hsp60* was amplified by PCR from pMV261.hyg using the primers Hsp60for-BstBI and Hsp60rev2-ClaI (Supplemental Table 2.1). The resulting PCR product was ligated into pJES113, which had been linearized with *Cla*I, to produce pJM157 (Fig. 2.1B). For construction of Library 2, genomic DNA was isolated from *M. tuberculosis* strain PM638 (Δ*blaC*). Digestion of genomic DNA and ligation into the single *Cla*I site upstream of '*blaC* in pJM157 was conducted as described above. After electroporation into *E. coli* DH5α (Invitrogen), approximately 8×10^5 CFU were pooled and used to isolate plasmid DNA (Qiagen).

Selection of exported 'BlaC fusions. Library DNA was electroporated into Δ*blaS M. smegmatis* strain PM759 (6, 18). The resulting transformation was plated onto 7H10 agar media without tween, containing 40 ug/ml lysine (Fischer Scientific), 50 μg/ml hygromycin and carbenicillin concentrations that ranged from 35-75 μg/ml and incubated at 37°C for a minimum of four days. The drug resistance of colonies that grew up on 7H10 agar with hygromycin and carbenicillin was confirmed by spot test analysis on 7H10 agar plates containing (i) both 50 μg/ml hygromycin and 45 μg/ml carbenicillin, and (ii) 50 μg/ml hygromycin only. Strains were further analyzed if spots revealed confluent growth on plates containing hygromycin and carbenicillin in comparison to the negative control strain (Δ*blaS M. smegmatis* with plasmid pJM113 that carries a promoter-less *'blaC*) (34).

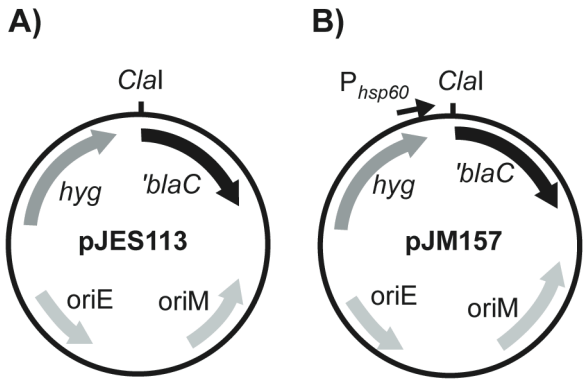

Figure 2.1. Plasmids used for construction of *M. tuberculosis*-'*blaC* fusion libraries. A) Library 1 was constructed by ligating fragments of *M. tuberculosis* H37Rv genomic DNA into the unique *Cla*I site located upstream of '*blaC* in pJES113. The truncated β-lactamase gene ('*blaC*) lacks a promoter, Shine-Dalgarno sequence, start codon, and signal sequence for export. (B) Library 2 was constructed with genomic DNA of *M. tuberculosis* PM638 (Δ*blaC*) ligated into the unique *Cla*I site located upstream of '*blaC* in pJM157. The pJM157 plasmid additionally has the mycobacterial *hsp60* promoter located upstream of the *Cla*I site, but lacks a Shine Dalgarno site and start codon for *'blaC*. Both library plasmids carry a selectable hygromycin-resistance gene (*hyg*) and have origins of replication for mycobacteria (oriM) and for *E. coli* (oriE).

Recovery of '*blaC* fusion plasmids. Plasmid DNA was transferred from carbenicillin and hygromycin-resistant Δ*blaS M. smegmatis* to *E. coli* DH5α by electroduction (2). A small amount of *M. smegmatis*

was transferred from a colony into 20 μl of ice-cold 10% glycerol. The suspension was mixed by vortexing, incubated on ice for 10 minutes, and added to 40 μl of electrocompetent *E. coli* DH5α. The mixture was transferred to a chilled 0.2 cm gap cuvette and pulsed using conditions typical for *E. coli* electroporation (25-μF, 200 Ohms, 2.5 kV). Immediately, 1 ml of LB broth was added to the cuvette, incubated at 37°C for 1 h and then plated on LB with hygromycin. Plasmid DNA was purified from hygromycin-resistant *E. coli* (Qiagen). The *M. tuberculosis* genomic DNA insert upstream of '*blaC* was identified by sequencing with the primer TnBlaCout (Supplemental Table 2.1).

Anti-BlaC antibody. A 6× histidine-tagged copy of *M. tuberculosis* BlaC was expressed from Y49 pTrcHisB-BlaC in *E. coli* DH5α (provided by Doug Kernodle), and purified by nickel affinity chromatography using HIS-Select Nickel Affinity Gel (Sigma) as described previously (59). Purified BlaC was eluted from the nickel column with 300 mM imidazole at a concentration of 1.3 mg/ml and used to immunize rabbits together with TiterMax Gold adjuvant (Sigma). Rabbit immunizations and polyclonal antisera collection was carried out by BioSource Custom Immunology Department (Hopkinton, MA).

Immunoblot analysis. Whole cell lysates of *M. smegmatis* and whole cell lysates of formalin-killed *M. tuberculosis* were prepared, as described previously (7, 19, 33), and analyzed by SDS-PAGE and immunoblot. Polyclonal BlaC antiserum was used in immunoblot analysis at a dilution of 1:10,000 and polyclonal 19kDa antiserum (provided by Douglas Young) was used at a dilution of 1:40,000. Anti-rabbit peroxidase-conjugated antibodies (Biorad) were used as secondary antibody to both anti-BlaC and anti-19kDa. Both monoclonal HA antiserum (Covance) and monoclonal GroEL HAT5/IT-64 antiserum (World Health Organization collection) were used at a dilution of 1:20,000, and anti-mouse peroxidase-conjugated antibodies were used as secondary antibodies (Biorad).

Construction of specific *M. tuberculosis* -'BlaC fusions

The nucleotides encoding the signal sequences of *Rv2041c* and *Rv2525c* were PCR amplified using the primers rv2041ssF, 2401ssR or 2525F and 2525R, respectively (Supplemental Table 2.1). Both the forward and reverse primers encoded a *BstB*I site, and the forward primers also include the Shine-Dalgarno sequence from *M. tuberculosis hsp60*. The amplified fragments were ligated into pCR2.1 (Invitrogen), digested with *BstB*I and then ligated into *Cla*I digested pJM157, upstream of *'blaC.*

Construction of expression constructs for HA-tagged *M. tuberculosis* proteins.

(i) *hsp60* promoter-driven HA fusions: Rv0774c-HA, ΔssRv0774c-HA, Rv2843-HA, ΔssPlcB-HA. Oligonucleotide primers were designed to amplify full-length or truncated genes from *M. tuberculosis* genomic DNA. Forward and reverse primers were designed each with 5' extension sequences carrying *Msc*I and *Hin*dIII restriction sites, respectively (Supplemental Table 2.1). The resulting PCR product was first cloned into the pCR2.1 (Invitrogen) and sequenced. The cloned gene was then digested with *Msc*I and *Hin*dIII and the appropriate fragment was isolated and ligated into the mycobacterial shuttle vector pJSC77 (21), digested with *Msc*I and *Hin*dIII, which carries the *hsp60* promoter and multiple cloning site upstream of a C-terminal HA tag (Table 2.2). Due to a *Msc*I site within the *Rv2843* gene, the strategy was revised in this instance and a *Nru*I site was included instead of *Msc*I on the forward primer (Supplemental Table 2.1).

(ii) native promoter-driven HA fusions: PlcB-HA, Rv0315-HA. Oligonucleotide primers were designed to amplify a fragment of *M. tuberculosis* genomic DNA which included the full-length gene of interest and upstream sequence containing the putative native promoter. Forward primers were designed each with 5' extension sequences carrying *Xba*I and *Hin*dIII restriction sites, respectively (Supplemental Table 2.1). The resulting PCR product was first cloned into the pCR2.1 (Invitrogen) and sequenced. The cloned gene and promoter was then digested with *Xba*I and *Hin*dIII and the appropriate fragment

was isolated and ligated into the mycobacterial shuttle vector pJSC77, digested with *Xba*I and *Hin*dIII, which removes the *hsp60* promoter (Table 2.2).

(iii) **PlcB(KK)-HA**. The construct in which the codons for the twin arginine pair of PlcB were mutated to encode lysines (KK) was generated as follows. The PCR primers plcBKKfor2 and plcBKKrev3 overlapped at the site of mutation and were used in inverse PCR to generate a product from pJM199. The resulting PCR product was then end-repaired using the EndIt Kit (Epicentre) following the manufacturer's instructions and self-ligated to create pJM216.

Subcellular fractionation. Cell wall, membrane, and soluble fractions were prepared by differential ultracentrifugation as described previously (19, 44). Briefly, 100 ml cultures of *M. tuberculosis* were harvested by centrifugation at 3,000 x *g*. Cell pellets were sterilized by gamma-irradiation in a JL Shephard Mark I 137Cs irradiator (Department of Radiobiology, University of North Carolina at Chapel Hill) with a dose of 2.4 megarads. All subsequent steps were performed at 4°C. Pellets were resuspended in 4 ml of breaking buffer (PBS, 1mM phenylmethylsulfonyl fluoride, 0.6 μg/ml each of DNase and RNase, and a cocktail of protease inhibitors [2 μg/ml each of aprotinin, E-64, leupeptin, and pepstatin A and 100 μg/ml Pefabloc SC]) and then lysed in a French pressure cell. Unbroken cells were pelleted at 3,000 x *g* for 20 min to generate a clarified whole-cell lysate, which was centrifuged at 27,000 x *g* for 30 min to pellet the cell wall. The supernatant was centrifuged at 100,000 x *g* for 2 h to separate the membrane fraction from the soluble fraction. The cell wall and membrane fractions were washed once and then resuspended in PBS.

Bioinformatic identification of putative twin-arginine signal sequences. Protein sequences corresponding to the 4056 predicted open reading frames (ORFs) of *M. tuberculosis* H37Rv were obtained from TubercuList (Institut Pasteur [http://genolist.pasteur.fr/TubercuList/]) and were entered as query sequences in the TATFIND v.1.4 (45) (http://signalfind.org/tatfind.html) and TatP v.1.0 (4)

(http://www.cbs.dtu.dk/services/TatP/) search algorithms. The output for TATFIND v.1.4 is either "True" or "False" for a given peptide. TatP v.1.0 has multiple outputs. We used the default search criteria RRx[FGAVML][LITMVF] and selected only those proteins that had both a predicted tri-partite signal peptide with Tat motif according to the prediction program. We also reviewed the list of predicted Tat exported proteins of *M. tuberculosis* defined by the TigrFAM motif (TIGR01409) (52). This list was obtained directly from the TIGR website (http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?database=ntmt02).

RESULTS

Selection of exported *M. tuberculosis* ORF-'BlaC fusions

A truncated *M. tuberculosis* β-lactamase 'BlaC, lacking its native signal sequence, can be used in mycobacteria as a reporter of export (34)**.** A special feature of the 'BlaC reporter is that it only works when exported by the Tat pathway. In a β-lactam-sensitive background, such as *M. smegmatis* Δ*blaS*, exported 'BlaC fusion proteins can be detected by their ability to promote growth in the presence of the β-lactam antibiotic carbenicillin.

To experimentally identify *M. tuberculosis* ORFs with functional Tat signal sequences, we constructed genomic *M. tuberculosis* libraries upstream of the truncated *'blaC* reporter. Two libraries were constructed in multicopy vectors, pJES113 and pJM157, each of which carries truncated *'blaC* immediately downstream of a unique *Cla*I cloning site (Fig. 2.1). The difference between the vectors is that pJM157 contains the mycobacterial *hsp60* promoter upstream of the *Cla*I site to drive expression from genomic fragments that lack a promoter. The *hsp60* sequence in pJM157, however, does not include a Shine-Dalgarno site or start codon; these elements must be provided by the genomic insert. For Library 1, *M. tuberculosis* genomic DNA was prepared from wild-type strain H37Rv and for Library 2, the genomic DNA was prepared from the Δ*blaC* mutant of *M. tuberculosis*. In both cases the genomic DNA was cut with *Cla*I-compatible endonucleases for ligation into the vectors.

The libraries were electroporated into the Δ*blaS* mutant of *M. smegmatis* and directly plated on 7H10 medium containing carbenicillin. Plasmids expressing exported fusion proteins were selected by their ability to promote growth in the presence of carbenicillin. For Library 1, 101 carbenicillin-resistant colonies were obtained from an estimated 1×10^6 *M. smegmatis* transformants. For Library 2, 29 carbenicillin-resistant colonies were obtained from an estimated 9×10^5 transformants. Following confirmation of carbenicillin resistance of individual transformants, plasmid DNA was isolated and sequenced to determine the identity of the genomic DNA insert. With one notable exception, all plasmids sequenced revealed an in-frame fusion with *'blaC*. The exception was plasmids where the full-length *M. tuberculosis blaC* gene was cloned. In these cases, the *blaC* insert did not need to be cloned in-frame with the reporter sequence on the plasmid. BlaC was identified in 50% of the carbenicillin resistant clones from Library 1. In Library 2 this problem was avoided by using genomic DNA from Δ*blaC M. tuberculosis*.

From the two libraries, we identified amino-terminal sequences of ten unique *M. tuberculosis* proteins that promote export of the 'BlaC reporter (Table 2.3). To confirm that the fusion proteins identified were exported in a Tat-dependent manner, we rescued the fusion plasmids and electroporated them into a double Δ*blaS* Δ*tatA M. smegmatis* mutant. This allowed us to test for export in the absence of a functional Tat pathway. All ten fusions that conferred carbenicillin resistance in a Δ*blaS* mutant background failed to confer carbenicillin resistance in the Δ*blaS* Δ*tatA* strain. This indicated that all the fusion proteins identified require the Tat pathway to export functional 'BlaC.

Direct testing of candidate *M. tuberculosis* Tat signal sequences

The *M. tuberculosis* sequences fused to 'BlaC in the ten active fusions were all predicted to contain signal sequences, according to the Signal P 3.0 prediction algorithm (Fig. 2.2A) (3). Evaluation of the exported fusions revealed that the junction with 'BlaC always occurred close to the predicted signal sequence cleavage site of the *M. tuberculosis* protein. The greatest distance we observed between a

predicted cleavage site and the ‘BlaC fusion junction was 34 amino acids. This revealed a requirement for an appropriately positioned restriction enzyme site near the cleavage site in order to identify a Tat signal sequence in our libraries. It also suggested some proteins may have been missed for this reason. For example, in previous work we showed the PlcA signal sequence promotes Tat export of the ‘BlaC reporter (34), but PlcA was not identified in the libraries. From genome gazing and bioinformatic predictions (discussed below), we selected Rv2041c and Rv2525c as candidate Tat substrates that may have been missed. Construction and direct testing of ssRv2041c-‘BlaC and ssRv2525c-‘BlaC fusion proteins revealed that both confer resistance to β-lactam in a Tat-dependent manner. This provides experimental validation of these additional *M. tuberculosis* Tat signal sequences (Fig. 2.2B).

Table 2.3. Functional Tat signal sequences in *M. tuberculosis* proteins identified with the ‘BlaC reporter.

Protein	**Description**	**No. Times Identified in Libraries**	***In slico* predictions**		
			TatP	**TAT FIND**	**Tigr FAM**
Rv0063	Possible oxidoreductase	1	…	…	×
Rv0315	Possible beta-1,3-glucanase precursor	41	×	×	×
Rv0483 (LprQ)	Possible conserved lipoprotein	2	…	…	…
Rv0519c	Possible lipase	5	×	…	×
Rv0774c	Possible lipase; short-chain alcohol dehydrogenase family	12	×	×	×
Rv0846c	Possible multicopper oxidase	1	×	×	×
Rv2068c (BlaC)	Class A beta-lactamase	63	×	×	×
Rv2350c (PlcB)	Phospolipase C	1	×	×	×
Rv2833c (UgpB)	Possible glycerol-3-phosphate-binding lipoprotein	3	×	×	×
Rv2843	Possible conserved transmembrane Ala-rich protein	1	…	×	×
Rv2041c	Possible sugar-binding lipoprotein	0*	×	×	×
Rv2351c (PlcA)	Phospolipase C	0*	×	…	×
Rv2525c	Conserved hypothetical protein	0*	×	×	×

*Demonstrated by direct testing only

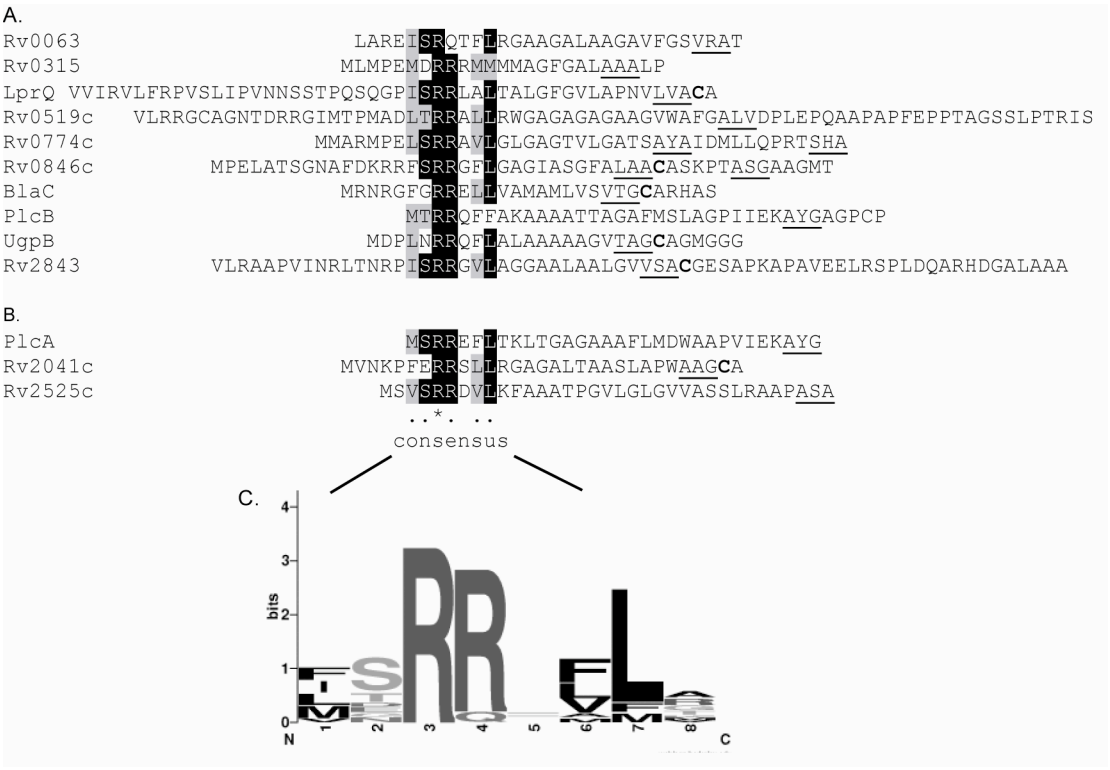

Figure 2.2. ***M. tuberculosis*** **Tat signal sequences capable of driving the export of 'BlaC.** Multiple sequence alignment (PROMALS, http://prodata.swmed.edu/promals/promals.php) of *M. tuberculosis* sequences capable of directing the export of functional 'BlaC. **(A)** Alignment of sequences identified from the genomic libraries. For a given protein, the shortest stretch of amino acids up to the fusion junction with 'BlaC is shown. **(B)** Alignment of sequences from fusions directly tested. Identical amino acids are shaded in black, and similar amino acids are shaded in gray. Underlined are signal sequence cleavage sites predicted by Signal P (3). Cysteine residues in bold indicate conserved residues in lipobox motifs that are the predicted site of lipid modification for lipoproteins. **(C)** Graphical representation of the sequence alignment of the consensus region of the 13 twin-arginine signal sequences shown to export functional 'BlaC. The height of each stack represents the degree of sequence conservation and the size of each letter is proportional to the frequency of the corresponding amino acid in that position (http://weblogo.berkeley.edu).

Alignment of functional Tat signal sequences identified with the 'BlaC reporter.

From studies largely conducted in *E. coli*, the generally accepted twin-arginine consensus motif is S/T-R-R-x-F-L-K (x is any polar amino acid). The twin arginines are nearly always invariant and the frequency of occurrence of the other amino acids is reported to exceed 50% (5, 31, 39). Amino acid alignment of the *M. tuberculosis* signal sequences we identified revealed that all but one contained the twin-arginine dipeptide (Fig. 2.2). The exception was the Rv0063 signal sequence, which has a glutamine in the position where the second arginine would be (R-Q-T-F-L). In terms of the other amino acids in the consensus, all were present in $\geq$ 40% of the sequences except for the final lysine (K). The alignment also revealed conservation of a hydrophobic residue just prior to the S/T amino acid in the consensus.

Comparison to *in silico* predicted Tat signal sequences.

Multiple Tat signal prediction programs exist, but their ability to accurately and comprehensively identify Tat substrates within mycobacteria is unresolved. We applied two of these web-based programs, TatP v1.0 (4) and TATFIND v1.4 (45) to the *M. tuberculosis* H37Rv genome sequence and compared the output to Tat signal sequences predicted by a third prediction program devised by The Institute for Genomic Research (TIGR) (TIGR01409) (52). Of the 4056 ORFs in the *M. tuberculosis* H37Rv genome (11), 95 were predicted to encode proteins with Tat signal sequences by at least one of the prediction programs (Supplemental Table 2.2). There is surprisingly limited overlap between the algorithms, with only eleven proteins being predicted by all three programs (Fig. 2.3A).

We next compared the signal sequences we identified experimentally (Table 2.3, Fig. 2.2) to the *in silico* predictions of the annotated *M. tuberculosis* genome. Eight of the proteins we identified were predicted by all three programs and three were predicted by two programs (Fig. 2.3B). Of the remaining two signal sequences, one was Rv0063, which lacks the twin arginine and was only identified by the TigrFAM prediction program, and the other was LprQ, which was not predicted by any program.

Assessment of Tat dependence for full-length proteins with functional Tat signal sequences

In addition to there being a requirement for a Tat signal sequence, the mature domain also plays a role in whether a protein is a Tat substrate. For this reason, we tested full-length versions of a subset of the proteins we identified to see if they were Tat-dependent.

Signal sequence cleavage is commonly used as an indicator of protein export (37). To establish the utility of this approach for monitoring export of *M. tuberculosis* Tat substrates expressed in *M. smegmatis*, we first assayed signal sequence cleavage of full length BlaC. Polyclonal antibodies were raised against BlaC and used to detect BlaC expressed in whole cell lysates of Δ*blaS M. smegmatis* by immunoblot (Fig. 2.4). In the presence of a functional Tat apparatus, BlaC is observed as a predominant band running at 30 kDa, which is the predicted size of the exported and processed mature species. Expression of BlaC(KK), which has a KK substitution for the 'RR' dipeptide, produced a slower migrating species which is consistent with a lack of export and accumulation of full-length unprocessed BlaC precursor (Fig. 2.4). In the absence of a functional Tat pathway (in a Δ*blaS* Δ*tatA* mutant), a larger presumptive precursor species was observed although some smaller mature protein was detected as well. These experiments indicated that the protein species observed for each of these strains was a good indicator of Tat- and twin-arginine-dependent export.

We then tested full-length versions of four additional proteins identified with the 'BlaC reporter in wild-type and Δ*tatC M. smegmatis*. Proteins were expressed as a C-terminal fusion to the HA epitope tag and immunoblot analysis was performed on whole cell lysates. For Rv0315 and Rv2843, Tat-dependent processing was observed in *M. smegmatis* which is indicative of the full-length proteins being Tat-exported. A smaller molecular weight and presumably processed form of the protein was seen when expressed in wild-type, but this species was significantly reduced in abundance in the Δ*tatC* mutant (Fig. 2.5A). Three bands were observed for Rv0315-HA expressed in wild-type *M. smegmatis*. The largest

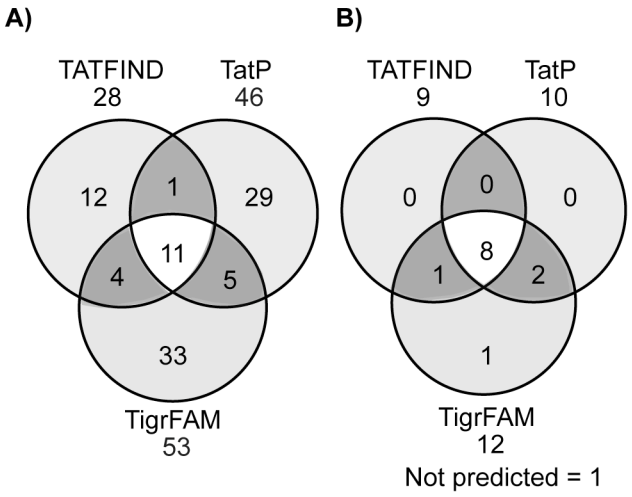

Figure 2.3. Distribution of predicted and experimentally verified *M. tuberculosis* Tat signal sequences. (A) Venn diagram indicating the distribution of *M. tuberculosis* proteins predicted to have Tat signal sequences by TATFIND v.1.4, TatP v.1.0, and/or TigrFAM (TIGR01409) algorithms. **(B)** Venn diagram comparing the proteins predicted *in silico* with the *M. tuberculosis* Tat signal sequences experimentally identified with the 'BlaC reporter. Rv0063 was predicted by TigrFAM only. LprQ was not predicted by any of the programs.

molecular weight band is similar in size (~34 kDa) to that of the predicted Rv0315 precursor and is present in both wild-type and Δ*tatC*. The two smaller protein species are absent or greatly reduced in the Δ*tatC* mutant background. This suggests that Rv0315 is subject to multiple processing events.

For the other two full-length proteins tested, no obvious Tat-dependence was observed in immunoblots of *M. smegmatis* whole cell lysates (Fig. 2.5B). One of these proteins was the virulence factor PlcB-HA. The protein species seen in the wild-type and Δ*tatC M. smegmatis* strains migrated at the predicted precursor size of 57 kDa. Moreover, when compared to a truncated form of PlcB-HA lacking the predicted signal sequence (ΔssPlcB-HA) the full-length product we expressed migrated slower than the expected mature product (Fig. 2.5B). Similar results were obtained with Rv0774c-HA and a truncated ΔssRv0774c-HA protein expressed in *M. smegmatis*. This suggested that these two predicted Tat substrates were not being processed or exported when expressed in *M. smegmatis*.

Full-length PlcB is exported by the Tat pathway when expressed in its native host *M. tuberculosis*

Since *M. smegmatis* does not have phospholipase C homologs, we considered the possibility that PlcB-HA can only be exported by its native host *M. tuberculosis*. To test this idea, we expressed full length PlcB-HA in *M. tuberculosis* H37Rv and assayed whole cell lysates by immunoblot analysis. Unlike what was observed in *M. smegmatis*, immunoblots for PlcB-HA in *M. tuberculosis* whole cell lysates yielded two products: a larger species that migrated like the full length precursor seen in *M. smegmatis* and a smaller species that migrated like the expected mature ΔssPlcB-HA product (Fig. 2.6A). This suggested that in *M. tuberculosis*, PlcB-HA is exported and processed. Subcellular fractions prepared from the *M. tuberculosis* strain were used to show that the observed faster migrating product is exported. Soluble, membrane, and cell wall fractions were prepared from whole cell lysates

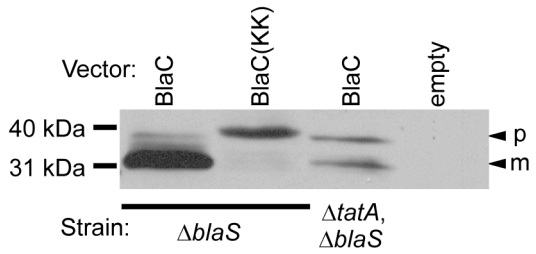

Figure 2.4. Tat-dependent processing of BlaC in *M. smegmatis*. Full-length BlaC was assessed for signal sequence processing by immunoblot analysis of whole cell lysates (WCL). Full length BlaC was expressed in a *tat*+ strain (Δ*blaS*) or a *tat*- deletion strain (Δ*blaS* Δ*tatA*) of *M. smegmatis*. BlaC[KK] was expressed in a *tat*+ strain. WCL of these strains were separated by SDS PAGE and compared to a WCL of Δ*blaS M. smegmatis* carrying an empty vector (empty). Bands were detected by immunoblot analysis using anti-BlaC antibodies. The presumed precursor (p) and mature (m) forms of select proteins are indicated by arrowheads.

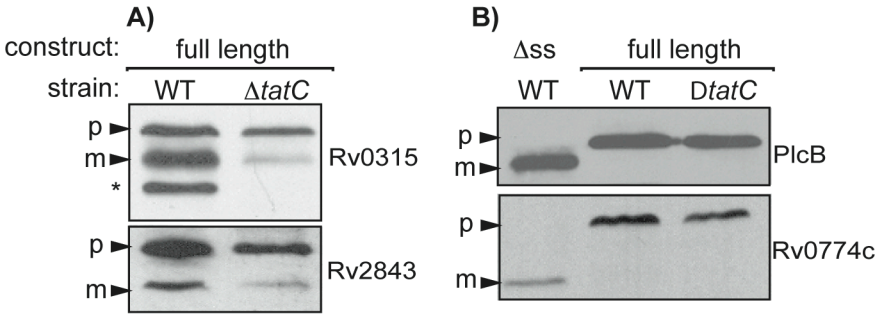

Figure 2.5. Assessment of signal sequence processing of *M. tuberculosis* full-length candidate Tat substrates. (A) Full-length constructs expressing HA-tagged Rv0315 and Rv2843 were assessed for signal sequence processing in WT or Δ*tatC M. smegmatis*. Whole cell lysates were prepared and proteins were separated by SDS-PAGE. Bands were detected by immunoblot analysis using anti-HA antibodies. The presumed precursor (p) and mature (m) forms of select proteins are indicated by arrowheads. The asterisk indicates a second proteolytic cleavage product for Rv0315. (B) Constructs of HA-tagged PlcB and Rv0774c lacking their presumed signal sequences (Δss) were run alongside full-length versions PlcB-HA and Rv0774-HA expressed in WT or Δ*tatC M. smegmatis.*

of H37Rv expressing PlcB-HA and analyzed by immunoblot (Fig 2.6B). Of the two protein species seen in the whole cell lysate, the larger product, the presumptive non-exported precursor, was in the

soluble cytosolic fraction and the smaller product was primarily in the cell wall. Thus, when expressed in *M. tuberculosis* the PlcB-HA protein is processed and exported to the cell wall. As controls for the fractionation, we showed the GroEL protein and the 19kDa lipoprotein were enriched in the soluble and cell envelope (cell wall and membrane) fractions, respectively.

The essential nature of the Tat pathway in *M. tuberculosis* precludes testing full-length PlcB-HA in a *M. tuberculosis* Δ*tat* mutant. To address whether full-length *M. tuberculosis* PlcB is a Tat substrate, the 'RR' pair in the PlcB signal sequence was replaced with 'KK'. When the PlcB(KK)-HA protein was expressed in *M. tuberculosis* a single precursor sized species was observed (Fig 2.6C). Together, these experiments demonstrated that PlcB is a Tat substrate exported to the cell wall in *M. tuberculosis* but not in *M. smegmatis*.

DISCUSSION

The Tat pathway is important to the virulence and physiology of several bacterial pathogens. Because it is absent in mammals, it has been considered a target for development of new antibacterial agents (8, 31). Here we used 'BlaC as a reporter to identify proteins with functional Tat signal sequences and begin understanding the role Tat export plays in *M. tuberculosis*. This approach allows for the experimental investigation of Tat export in a system where obtaining a Δ*tat* mutant is not feasible. 'BlaC is not the only recognized Tat specific reporter (14, 63), but it is the only such reporter that works in a direct selection, as opposed to a screen. This property of 'BlaC allowed us to exploit it on a genome-wide level. Previously, Tat reporters have only been used to validate pre-selected candidates (49, 57, 58, 63).

All active 'BlaC fusions obtained from our libraries were in-frame with an ORF, had predicted signal sequences, and were confirmed to be exported in a Tat-dependent manner. This attests to the

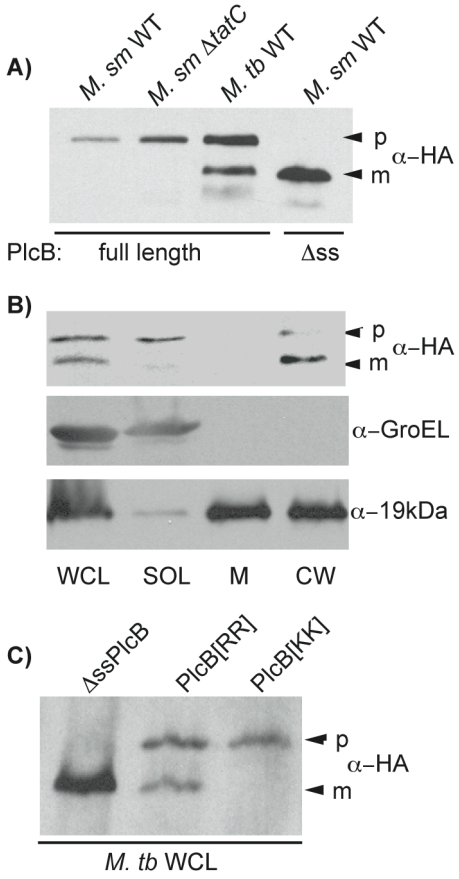

Figure 2.6. Full-length PlcB is processed and exported in *M. tuberculosis* but not in *M. smegmatis*. **(A)** Full length PlcB-HA and PlcB-HA lacking its signal sequence (Δss) were expressed in wild-type (*M. sm* WT) and *tat*- deletion (*M. sm* Δ*tatC*) strains of *M. smegmatis*, or a wild-type strain of *M. tuberculosis* (*M. tb* WT). Bands were detected by immunoblot analysis using anti-HA antibodies. The presumed precursor (p) and mature (m) forms of PlcB are indicated by arrowheads. **(B)**. Irradiated cultures of *M. tuberculosis* expressing HA-tagged full length PlcB-HA were fractionated from whole cell lysates (WCL) into soluble (S), membrane (M) and cell wall (CW) fractions. Immunoblot analysis was performed to localize HA-tagged PlcB-HA and GroEL and 19kDa lipoprotein controls. **(C)** HA tagged truncated ΔssPlcB, full length PlcB[RR] (wild-type), and PlcB[KK], were expressed in *M. tuberculosis*. WCL of each strain were analyzed by immunoblot analysis.

power of the 'BlaC reporter. Our objective was to comprehensively identify the Tat signal sequences of *M. tuberculosis*. Our two libraries were composed of ~2×10^6 plasmids combined, which is sufficiently

complex to have >99% probability of every *M. tuberculosis* ORF being represented as a single in-frame '*blaC* fusion (48). However, the unanticipated requirement for a fusion junction to be positioned just beyond the signal sequence cleavage site means that, despite this level of complexity, some proteins were not picked up. This restriction most likely reflects extended sequences of a mature protein preventing 'BlaC folding into an active conformation compatible with Tat export. Future libraries could be improved by increasing the complexity and using smaller randomly sheared genomic DNA inserts to overcome limitations that include the lack of available restriction enzyme cleavage sites.

Of the thirteen Tat signal sequences identified with the reporter, many are in proteins (Rv0063, Rv0315, Rv0774c, Rv2525c, BlaC, and PlcB) previously shown to be secreted or cell wall-associated by proteomics (22, 24, 32, 43, 51). Six of them (Rv0846c, Rv2041c, Rv2843, BlaC, LprQ, and UgpB) are predicted lipoproteins with a lipobox motif in their signal sequence (Fig 2.2A), which predicts cell envelope localization (56). Recently, in the archaea *Haloferax volcanii* it was shown that lipoproteins can be Tat substrates (20). BlaC, PlcA, and PlcB are the only proteins we identified that have demonstrated functions (β-lactamase and phospholipase C activities, respectively) (29, 43, 59). The Plc proteins are also shown to function in *M. tuberculosis* pathogenesis. Simultaneous deletion of multiple *plc* genes results in attenuated virulence of *M. tuberculosis* in a mouse model of infection (43).

Much less is known about the remaining Tat signal sequence-containing proteins we identified (Table 2.3). Sequence analysis suggests these proteins have diverse functions as lipases (Rv0519c and Rv0774c), a copper oxidase (Rv0846c), a glycosyl hydrolase (Rv0315), an oxidoreductase (Rv0063), and substrate-binding proteins of sugar uptakes systems (UgpB and Rv2041c). Rv2525c is one of the proteins identified with no predicted function. On the basis of bioinformatic predictions this protein was previously hypothesized to be a *M. tuberculosis* Tat substrate and, because of co-regulation with other genes, was suggested to have a role in cell wall biogenesis (47). Both UgpB and Rv2041c are predicted by transposon saturation mutagenesis (TraSH) to be essential to *M. tuberculosis* (50), providing a potential clue as to why the Tat pathway can not be inactivated in *M. tuberculosis*.

An advantage of using a genetic reporter to directly identify functional Tat signal sequences is that there is no imposed requirement that conserved sequence elements, as defined by other studies, need be present. In this regard, the functional Tat signal sequences we identified should be useful for refining Tat prediction algorithms. Most of the sequences we experimentally identified were predicted by at least two of three Tat signal sequence prediction programs consulted. This suggests that the common core elements of these programs are the best predictors. The signal sequence of Rv0063, which possesses a RQ in the Tat motif, was the only one of our experimentally identified sequences lacking a twin-arginine dipeptide. Our identification of Rv0063 is consistent with the recent demonstration that its signal sequence can direct Tat export of an agarase reporter when expressed in *Streptomyces lividans* (63). Although no natural Tat substrate is known to have a RQ pair, there are two examples of bacterial Tat signal sequences lacking a 'RR' dipeptide (26, 27). In addition, a RQ substitution in the Tat signal sequence of the *E. coli* TorA protein is able to promote Tat export of a GFP reporter (14). Interestingly, the TigrFAM program predicted Rv0063 to have a Tat signal sequence (52). We also identified one protein, LprQ, which was not identified by any of the three prediction programs. LprQ was likely missed due to its extended N-terminus upstream of the twin-arginine motif (Fig 2.2A). However, it is also possible that the true start codon of LprQ is incorrectly annotated, as there are additional GTG codons between the annotated start and the twin-arginine motif.

In addition to having an amino-terminal Tat signal sequence, there are features of the mature domain of a protein, which must be folded, that determine whether it can be translocated by the Tat pathway (15). Recently, some putative Tat signal sequences of *E. coli* were shown to be able to promote export of a fused reporter through Sec or Tat pathways, depending on the unfolded or folded nature of the reporter (58). These apparently promiscuous signal sequences highlight the importance played by the mature domain of a protein. The basic question of how often it is that a functional Tat signal sequence is present on a bona fide Tat substrate has only begun to be investigated (58). For this reason, we examined full-length versions of a subset of proteins we identified in wild-type and Δ*tat M.*

smegmatis. Three proteins (BlaC, Rv0315, and Rv2843) showed a Tat-dependent effect, which indicates that the native proteins are subject to Tat export. For Rv0315-HA, the immunoblot analysis revealed three protein species in whole cell lysate of wild-type *M. smegmatis* and two species in lysate of *tat* mutant *M. smegmatis* (Fig. 2.5A). A homologous β-glucanase of *Bacillus circulans* is subject to progressive proteolytic processing post-export (61). Similar proteolytic processing of Rv0315 would explain the multiple species seen by immunoblot.

For two other *M. tuberculosis* proteins tested there was no difference in the protein species seen in whole cell lysate from wild-type and Δ*tat M. smegmatis* (Fig. 2.5B). In these cases, the size of the full-length protein species observed was larger than that of the predicted mature protein. Thus, no signal sequence cleavage or obvious export occurred when full-length versions of these proteins were expressed in *M. smegmatis*.

The fact that PlcB-HA was unaffected in a *M. smegmatis tat* mutant was surprising since the Plc proteins are demonstrated to be exported to cell wall fractions of *M. tuberculosis* (43). Moreover, virtually all phospholipase C orthologs in a wide-variety of organisms have predicted Tat signal sequences (16) and there are bacterial Plc enzymes proven to be exported by the Tat pathway (9, 36, 46, 60). In contrast to what was seen in *M. smegmatis*, when expressed in *M. tuberculosis* two species of PlcB-HA were evident by immunoblot analysis. The faster of the two species was exported to the cell wall and migrated at the same molecular weight as ΔssPlcB-HA, the expected size of processed mature PlcB. A KK substitution of the twin-arginine motif of full-length PlcB prevented its processing in *M. tuberculosis*. These results demonstrate that in its native host PlcB is an authentic twin-arginine dependent protein. These data provide an important link between the Tat pathway and virulence factor export in *M. tuberculosis*.

The above result is also interesting in suggesting that Tat export of a *M. tuberculosis* virulence factor requires a pathogen-specific component. There are a small number of Tat substrates shown to require dedicated chaperones for export (28, 38, 39). There are also examples of chaperones, such as

DnaK, that work with a broader collection of Tat substrates (23, 40). Perhaps the pathogen-specific factor is a chaperone for PlcB that is present in *M. tuberculosis* but lacking in *M. smegmatis*. Although the exact function of these chaperones remains to be discerned, proposals for how they work include promoting proper folding, protecting the Tat substrate from degradation or delivery to the translocon before folding is complete, and delivering the substrate to the translocase (39). Some of these functions relate to the mature domain of the protein, which could account for why a ssPlcB-'BlaC protein was exported by *M. smegmatis* but the full-length PlcB-HA was not. Notably, PlcH of *P. aeruginosa* requires PlcR chaperones for its export (12). However, no obvious PlcR homologues exist in *M. tuberculosis*. Another possibility is that a conserved component of the *M. smegmatis* and *M. tuberculosis* Tat systems differs in its substrate recognition abilities.

The work presented here demonstrates the power of using a Tat-specific reporter to identify functional Tat signal sequences without any preconceived bias regarding the features that define them. By examining full-length versions of proteins in wild-type and Δ*tat M. smegmatis*, we showed three of the proteins identified to be true Tat substrates. The results with the other two proteins tested emphasize the importance of domains beyond the Tat signal sequence and bacteria-specific factors in defining a true Tat substrate. In our quest to prove the virulence factor PlcB is a true Tat substrate, we discovered this protein is exported by the Tat pathway only its native host, indicating the existence and requirement for pathogen-specific host factors in its Tat export.

Acknowledgements

This research was funded by a grant from NIAID (AI54540 to M.B) and a developmental grant from the University of North Carolina at Chapel Hill Center for AIDS Research (NIH #9P30 AI50410-

04). J.A.M. was also supported by the Training in Sexually Transmitted Diseases and AIDS NIH NIAID training grant 5T32 AI07001-28 and a Society of Fellows dissertation completion fellowship from the Graduate School at the University of North Carolina at Chapel Hill. J.R.M was also supported by an NIH Cell and Molecular Biology training grant (NIH 5-T32- GM008581). J.S.S. was supported by NSF REU grant DBI-0453235. We would like to thank Doug Young for 19kDa lipoprotein antibodies, Doug Kernodle for providing us with the HIS-tagged BlaC expression vector, Adrienne Cox and Henry Gibbons for assistance with irradiation and fractionation of *M. tuberculosis* cultures, and Ruth Silversmith for advice on purification of HIS-tagged BlaC. We also thank Martin Pavelka and members of the Braunstein laboratory for critical review of the manuscript.

Attributions

J.A.M., J.R.M. and M.B. conceived of the experiments. J.A.M. and J.R.M. constructed the genomic libraries and conducted the majority of experiments. E.M.T. and J.S.S. assisted in screening the library for exported, Tat-dependent fusions. N.W.R. assisted in fractionating *M. tuberculosis* expressing recombinant fusions. This work has been previously published (*Journal of Bacteriology*; 2008 **190**:6248-38) and permission to reprint this work has been granted by the publisher.

REFERENCES

1. **Bange, F. C., F. M. Collins, and W. R. Jacobs, Jr.** 1999. Survival of mice infected with *Mycobacterium smegmatis* containing large DNA fragments from *Mycobacterium tuberculosis*. Tuber Lung Dis **79:**171-80.

2. **Baulard, A., C. Jourdan, A. Mercenier, and C. Locht.** 1992. Rapid mycobacterial plasmid analysis by electroduction between *Mycobacterium spp*. and *Escherichia coli*. Nucleic Acids Res **20:**4105.

3. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol **340:**783-95.

4. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

5. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? Mol Microbiol **22:**393-404.

6. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. *In* V. L. Clark and P. M. Bavoil (ed.), Methods in Enzymology, vol. 358. Academic Press, London.

7. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-90

.

8. **Bronstein, P., M. Marrichi, and M. P. DeLisa.** 2004. Dissecting the twin-arginine translocation pathway using genome-wide analysis. Res Microbiol **155:**803-10.

9. **Bronstein, P. A., M. Marrichi, S. Cartinhour, D. J. Schneider, and M. P. DeLisa.** 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae pv. tomato* DC3000 and its contribution to pathogenicity and fitness. J Bacteriol **187:**8450-61.

10. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. Mol Plant Microbe Interact **19:**200-12.

11. **Camus, J. C., M. J. Pryor, C. Medigue, and S. T. Cole.** 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. Microbiology **148:**2967-73.

12. **Cota-Gomez, A., A. I. Vasil, J. Kadurugamuwa, T. J. Beveridge, H. P. Schweizer, and M. L. Vasil.** 1997. PlcR1 and PlcR2 are putative calcium-binding proteins required for secretion of the hemolytic phospholipase C of *Pseudomonas aeruginosa*. Infect Immun **65:**2904-13.

13. **Cristobal, S., J. W. de Gier, H. Nielsen, and G. von Heijne.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. Embo J **18:**2982-90.

14. **DeLisa, M. P., P. Samuelson, T. Palmer, and G. Georgiou.** 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. J Biol Chem **277:**29825-31.

15. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-20.

16. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-83.

17. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. J Bacteriol **185:**760-71.

18. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. Microbiology **151:**521-32.

19. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-100.

20. **Gimenez, M. I., K. Dilks, and M. Pohlschroder.** 2007. *Haloferax volcanii* twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins. Mol Microbiol **66:**1597-606.

21. **Glickman, M. S., J. S. Cox, and W. R. Jacobs, Jr.** 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis.* Mol Cell **5:**717-27.

22. **Gomez, M., S. Johnson, and M. L. Gennaro.** 2000. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach. Infect Immun **68:**2323-7.

23. **Graubner, W., A. Schierhorn, and T. Bruser.** 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. J Biol Chem **282:**7116-24.

24. **Gu, S., J. Chen, K. M. Dobos, E. M. Bradbury, J. T. Belisle, and X. Chen.** 2003. Comprehensive proteomic profiling of the membrane constituents of a *Mycobacterium tuberculosis* strain. Mol Cell Proteomics **2:**1284-1296.

25. **Halbig, D., T. Wiegert, N. Blaudeck, R. Freudl, and G. A. Sprenger.** 1999. The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. Eur J Biochem **263:**543-51.

26. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. FEBS Lett **497:**45-9.

27. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. Biochem Biophys Res Commun **291:**146-9.

28. **Jack, R. L., G. Buchanan, A. Dubini, K. Hatzixanthis, T. Palmer, and F. Sargent.** 2004. Coordinating assembly and export of complex bacterial proteins. Embo J **23:**3962-72.

29. **Johansen, K. A., R. E. Gill, and M. L. Vasil.** 1996. Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. Infect Immun **64:**3259-66.

30. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. Infect Immun **74:**1768-76.

31. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. Annu Rev Microbiol **60:**373-395.

32. **Malen, H., F. S. Berven, K. E. Fladmark, and H. G. Wiker.** 2007. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. Proteomics **7:**1702-18.

33. **McCann, J. R., J. A. McDonough, M. S. Pavelka, and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. Microbiology **153:**3350-9.

34. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-79.

35. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. Trends Microbiol **9:**494-500.

36. **Ochsner, U. A., A. Snyder, A. I. Vasil, and M. L. Vasil.** 2002. Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. Proc Natl Acad Sci U S A **99:**8312-7.

37. **Oliver, D. B., and J. Beckwith.** 1981. *E. coli* mutant pleiotropically defective in the export of secreted proteins. Cell **25:**765-72.

38. **Oresnik, I. J., C. L. Ladner, and R. J. Turner.** 2001. Identification of a twin-arginine leader-binding protein. Mol Microbiol **40:**323-31.

39. **Palmer, T., F. Sargent, and B. C. Berks.** 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. Trends Microbiol **13:**175-80.

40. **Perez-Rodriguez, R., A. C. Fisher, J. D. Perlmutter, M. G. Hicks, A. Chanal, C. L. Santini, L. F. Wu, T. Palmer, and M. P. Delisa.** 2007. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. J Mol Biol.

41. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-40.

42. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of Enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71:**4908-4916.

43. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-17.

44. **Roberts, A. D., J. T. Belisle, A. M. Cooper, and I. M. Orme.** 1998. Specific models: Murine model of tuberculosis, p. 389-417. *In* S. H. Kaufmann and D. Kabeltz (ed.), Methods in Microbiology, vol. 25.

45. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-50.

46. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. Infect Immun **73:**2020-32.

47. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis,* increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-79.

48. **Sambrook, J., and D. W. Russell.** 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

49. **Santini, C. L., A. Bernadac, M. Zhang, A. Chanal, B. Ize, C. Blanco, and L. F. Wu.** 2001. Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. J Biol Chem **276:**8159-64.

50. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc Natl Acad Sci U S A **98:**12712-7.

51. **Schmidt, F., S. Donahoe, K. Hagens, J. Mattow, U. E. Schaible, S. H. Kaufmann, R. Aebersold, and P. R. Jungblut.** 2004. Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. Mol Cell Proteomics **3:**24-42.

52. **Selengut, J. D., D. H. Haft, T. Davidsen, A. Ganapathy, M. Gwinn-Giglio, W. C. Nelson, A. R. Richter, and O. White.** 2007. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. Nucleic Acids Res **35:**D260-4.

53. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol **4:**1911-9.

54. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. J Biol Chem **275:**11591-6.

55. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-60.

56. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-59.

57. **Thomas, J. D., R. A. Daniel, J. Errington, and C. Robinson.** 2001. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. Mol Microbiol **39:**47-53.

58. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. J Biol Chem.

59. **Voladri, R. K., D. L. Lakey, S. H. Hennigan, B. E. Menzies, K. M. Edwards, and D. S. Kernodle.** 1998. Recombinant expression and characterization of the major beta-lactamase of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother **42:**1375-81.

60. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. Embo J **20:**6735-41.

61. **Watanabe, T., N. Kasahara, K. Aida, and H. Tanaka.** 1992. Three N-terminal domains of beta-1,3-glucanase A1 are involved in binding to insoluble beta-1,3-glucan. J Bacteriol **174:**186-90.

62. **Wickner, W., and R. Schekman.** 2005. Protein translocation across biological membranes. Science **310:**1452-6.

63. **Widdick, D. A., R. T. Eijlander, J. M. van Dijl, O. P. Kuipers, and T. Palmer.** 2008. A facile reporter system for the experimental identification of twin-arginine translocation (Tat) signal peptides from all kingdoms of life. J Mol Biol **375:**595-603.

64. **WorldHealthOrganization** 2007, posting date. WHO Information tuberculosis fact sheet. http://www.who.int/mediacentre/factsheets/fs104/en/

Supplemental Table 2.1. Primers used in this study.

Target gene	Oligo name	Sequence (5' → 3')	Restriction enzyme site (5' extension)
Primers used for library construction			
blaC	Lib'blaCfor	ATCGATCCGGCATCGACAACCTTG	*Cla*I
	Lib'blaCrev	CCATGGCTATGCAAGCACACCGGCAAC	*Nco*I
hsp60	Hsp60for-BstBI	ATTCGAACGTGGCGAACTCCGTTGTAG	*Bst*BI
	Hsp60rev2-ClaI	TATCGATCGGGGATGAAACGGGGGTC	*Cla*I
Primers used for plasmid construction (HA-tagged constructs)			
plcB	PlcBMscIF	CGTGGCCATTACCCGCCGACAATTTTTTGC	*Msc*I
	PlcBMscF-2	CGTGGCCATTGGGCCTTGCCCCGGACATTTG	*Msc*I
	plcBMscF-3	GTCTAGAGTCCGGTGGCATTCGTGGTC	*Xba*I
	plcBXbaF-1	GTCTAGAACCCTTCCCGCTGCGATATCC	*Xba*I
	PlcBHindR	CGAAGCTTGACAGAGACCGCTGGGAATCC	*Hin*dIII
	plcBKKfor2	AAGCAATTTTTTGCTAAAGCAGCCGCCGCTAC	n/a
	plcBKKrev3	CTTGGTCATTCCGTCGACGGGGTGTTC	n/a
Rv0315	315MscF-2	GTCTAGAAGAACCAGCACGCCGACAAC	*Xba*I
	315HindR	CGAAGCTTGAAAGACGCGCACCCAGTCG	*Hin*dIII
Rv0774c	774MscF	CGTGGCCATTATGGCCCGCATGCCAGAGTTG	*Msc*I
	774HindR	CGAAGCTTGACCTGTGAGCAGCGGCGCGAAC	*Hin*dIII
	Xss774mscIF	CTGGCCAACGCACCGGCGGCGGCGATTGGCAC	*Msc*I
Rv2843	2843NruF	ATCGCGATTCTTAGAGCAGCACCAGTCATC	*Nru*I
	2843HindR	CGAAGCTTGTATCGACGGGCCTGAAGGCAC	*Hin*dIII
Primers used for making signal sequence-'*blaC* fusions			
Rv2525c	2525F	CTTCGAACCGGAGGAATCACTTCGCATGTCGG TCTCTCGGCGTGATG	*Bst*BI
	2525R	CTTCGAAGGCCGATGCCGGCGCGGCGCGCAAC G	*Bst*BI
Rv2041c	Rv2041ssF	CTTCGAACCGGAGGAATCATTCGCAATGATGG TCAAGCCGTTCGAGCGGCGCAG	*Bst*BI
	2041ssR	CTTCGAAAGCACACCCAGCGGCCCAGGGA	*Bst*BI
Primers used for sequence determination			
blaC	TnBlaCout	CAGAATGCGAACCGCTCATC	n/a

CHAPTER 3

β-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells

Jessica R. McCann[1], Justin A. McDonough[1], Martin S. Pavelka, Jr.[2], and Miriam Braunstein[1]

[1]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599-7290

[2]Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642

Mycobacterium tuberculosis is an intracellular pathogen that is able to avoid destruction by host immune defenses. Exported proteins of *M. tuberculosis*, which include proteins localized to the bacterial surface or secreted into the extracellular environment, are ideally situated to interact with host factors. As a result, these proteins are attractive candidates for virulence factors, drug targets, and vaccine components. Here we describe a new β-lactamase reporter system capable of identifying exported proteins of *M. tuberculosis* during growth in host cells. Because β-lactams target bacterial cell wall synthesis, β-lactamases must be exported beyond the cytoplasm to protect against these drugs. When used in protein fusions, β-lactamase can report on the subcellular location of another protein as measured by protection from β-lactam antibiotics. Here we demonstrate that a truncated TEM-1 β-lactamase lacking a signal sequence for export (‘BlaTEM-1) can be used in this manner directly in a mutant strain of *M. tuberculosis* lacking the major β-lactamase, BlaC. The ‘BlaTEM-1 reporter

conferred β-lactam resistance when fused to both Sec and Tat export signal sequences. We further demonstrate that β-lactamase fusion proteins report on protein export while *M. tuberculosis* is growing in THP-1 macrophage-like cells. This genetic system should facilitate the study of proteins exclusively exported in the host environment by intracellular *M. tuberculosis*.

INTRODUCTION

Tuberculosis is responsible for nearly two million deaths each year (47). *Mycobacterium tuberculosis*, the causative agent of this disease, is an intracellular pathogen and the ability of this bacterium to survive and grow in macrophages is essential to its virulence. Multiple processes are likely employed by *M. tuberculosis* to avoid destruction in macrophages. These include residing in a phagosome that fails to mature into an acidified phagolysosome and resisting reactive radicals (as reviewed in (32, 48). As in other bacterial pathogens, *M. tuberculosis* proteins exported beyond the cytoplasm to the bacterial cell envelope (comprised of the cytoplasmic membrane and cell wall) or secreted into the environment are ideally positioned to interact with host cell components and promote survival in macrophages. Consequently, exported and secreted proteins make good candidates for virulence factors, drug targets for disease intervention, and vaccine antigens.

Mycobacteria possess two conserved pathways for exporting proteins: the general secretion (Sec) pathway and the twin-arginine translocation (Tat) pathway (2, 21, 25, 28, 29, 35). These systems recognize precursor proteins synthesized with amino-terminal signal sequences and transport them across the cytoplasmic membrane (10, 26). The proteins exported by these pathways can remain associated with the cell envelope or be further secreted by the bacterium. The signal sequences of Sec and Tat substrates share a similar domain structure; however, Tat substrates are distinguished by the presence of the twin-arginine motif, R-R-x-ϕ-ϕ (ϕ = uncharged residue). The two pathways also differ in their mode of transport. Sec substrates are translocated across the cytoplasmic membrane in an

unfolded state, whereas Tat substrates are translocated in a folded conformation. *M. tuberculosis* also has at least two specialized protein export pathways: the SecA2-dependent system and the ESX-1 (ESAT-6) system (3, 15, 18, 30, 40). Interestingly, both pathways appear capable of secreting specific subsets of proteins that lack conventional Sec or Tat signal sequences.

In *M. tuberculosis*, proteomic and genetic methods have been used to experimentally identify proteins exported beyond the cytoplasm (reviewed in (21). The genetic methods rely on reporter enzymes that are fused to *M. tuberculosis* protein sequences and report on the subcellular location of the fusion proteins (4, 8, 12, 23, 46). Surrogate hosts such as non-pathogenic *Mycobacterium smegmatis* or *Escherichia coli* have been used in most of these studies, often because endogenous enzyme activities in *M. tuberculosis* precluded their use directly in the pathogen. The use of surrogate hosts is a problem for identifying proteins that are only exported by pathogenic *M. tuberculosis*.

β-lactamase is an export reporter that was not initially employed directly in *M. tuberculosis* because of endogenous β-lactam resistance. β-lactamase catalyzes the hydrolysis of β-lactams, a class of antibiotic that targets cell wall biosynthetic enzymes located outside of the cytoplasmic membrane. Therefore, β-lactamase must be exported beyond the cytoplasm to protect the bacterium from the drug. For this reason, when fused to another protein, it can be used as an export reporter with β-lactam resistance as a powerful indicator of export. We recently reported that a *ΔblaC* mutant of *M. tuberculosis*, lacking the chromosomally encoded β-lactamase BlaC, is β-lactam sensitive (14). Further, we showed that BlaC is a native Tat substrate and that a truncated ‘BlaC lacking a signal sequence can function as a reporter of Tat-dependent export directly in a *ΔblaC* mutant of *M. tuberculosis* (25). This was shown by fusing a Tat signal sequence to ‘BlaC and demonstrating that the resulting hybrid protein confers resistance to the β-lactam antibiotic carbenicillin in the *ΔblaC* background. Interestingly, the ‘BlaC reporter works with Tat but not Sec exported proteins. Here we expanded the β-lactamase tools that can be used directly in *M. tuberculosis* by demonstrating that the TEM-1 β-lactamase (BlaTEM-1), originally identified in a clinical isolate of *E. coli* (9), functions as an export reporter in the *ΔblaC*

mutant of *M. tuberculosis*. The 'BlaTEM-1 reporter has the significant advantage of being compatible with both Sec and Tat signal sequences.

The proteomic and genetic approaches used in previous work for identifying exported proteins of *M. tuberculosis* are limited by their reliance on *in vitro* grown bacteria. Consequently, a potentially interesting collection of proteins only exported or secreted while *M. tuberculosis* are inside host cells are missed. In this report, we demonstrate that β-lactamase reporters have the novel capability of identifying *M. tuberculosis* proteins that are exported during intracellular growth in β-lactam treated THP-1 macrophage-like cells. The system we describe will be of significant value for identifying the most interesting category of exported *M. tuberculosis* proteins – those exported during growth in the host environment.

MATERIALS AND METHODS

Bacterial strains, media and growth conditions

Escherichia coli DH5α was grown in Luria-Bertani medium (Fisher) supplemented with the following concentrations of antibiotics as required: carbenicillin, 100 μg/ml; kanamycin, 40 μg/ml. *M. tuberculosis* strains H37Rv (WT), PM638 (*ΔblaC*, H37Rv) (14) and all derivative strains were cultured in Middlebrook 7H9 medium or on Middlebrook 7H10 agar medium (Difco; BD Biosciences) supplemented with 10% ADS (0.5% BSA, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl), 0.5% glycerol, and 0.05% tween 80 (Fisher). Antibiotics for mycobacteria were used at the following concentrations: carbenicillin, 50 μg/ml; kanamycin, 20 μg/ml. 7H10 plates supplemented with carbenicillin lacked tween, as the combination of tween and carbenicillin appeared detrimental to growth of fusion-expressing strains.

Construction of '*blaTEM-1* fusion plasmids

Plasmids used in this study are listed in Table 3.1. All subcloned PCR products were sequenced and determined to be error free. Sequence encoding the mature domain (lacking the N-terminal signal

Table 3.1. Plasmids used in this study

Plasmid	Genotype	Description	Source
pCC1	*cat oriV ori2*	CopyControl (single copy) blunt cloning vector	Epicentre
pCR2.1	*bla aph* ColE1	TA cloning vector	Invitrogen
pMV261.*kan*	*aph* P_{hsp60} *oriM* ColE1	Multicopy mycobacterial shuttle plasmid	(41)
pMB219	*aph oriM* ColE1	Multicopy mycobacterial shuttle plasmid	This work
pMB222	*aph* P_{hsp60}-ss*plcB* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis plcB* signal sequence in pMV261 under control of *hsp60* promoter	(25)
pMB227	*aph* P_{hsp60}-ss*mpt63* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis mpt63* signal sequence in pMV261 under control of *hsp60* promoter	(25)
pMB228	*aph* P_{hsp60}-ss*mpt63-'blaC* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis ssmpt63-'blaC* in pMV261 under control of *hsp60* promoter	(25)
pJM109	*aph* P_{hsp60}-ss*fbpB-'blaC* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis ssfbpB-'blaC* in pMV261 under control of *hsp60* promoter	(25)
pJM111	*aph* P_{hsp60}-ss*plcB-'blaC* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis ssplcB-'blaC* in pMV261 under control of *hsp60* promoter	(25)
pJM113	*aph* P_{hsp60}-*'blaC* (*M. tuberculosis*) *oriM* ColE1	*M. tuberculosis 'blaC* (no signal sequence) in pMV261 under control of *hsp60* promoter	(25)
pJM114	*cat oriV ori2*	*E. coli 'blaTEM-1* cloned into pCC1	This work
pJES101	*aph* P_{hsp60}-ss*plcB-'blaTEM-1* (*E. coli*) *oriM* ColE1	*'blaTEM-1* from pJM114 cloned into pMB222	This work
pJES102	*aph* P_{hsp60}-*'blaTEM-1* (*E. coli*) *oriM* ColE1	*'blaTEM-1* from pJM114 cloned into pMV261	This work
pJES103	*aph* P_{hsp60}-ss*mpt63-'blaTEM-1* (*E. coli*) *oriM* ColE1	*'blaTEM-1* from pJM114 cloned into pMB227	This work
pJES125	*bla aph* ColE1	*M. tuberculosis mpt83* signal sequence and upstream sequence cloned into pCR2.1	This work
pJES128	*aph oriM* ColE1	*'blaTEM-1* from pJM114 cloned into *Bam*HI-linearized pMB219	This work
pJES129	*aph* P_{mpt83}-ss*mpt83-'blaTEM-1* (*E. coli*) *oriM* ColE1	*M. tuberculosis mpt83* signal sequence and upstream sequence from pJES125 cloned into pJES128	This work

sequence) of *E. coli* BlaTEM-1 was amplified from pUC19 plasmid DNA (Invitrogen) using the following primers: TEMbla1 (5'- AGATCTCACCCAGAAACGCTGGTGAAAG) and TEMbla2 (5'-GTTACCAATGCTTAATCAGTGAGGCACC). The resulting PCR product was cloned into the pCC1 vector (Epicentre) to generate pJM114. The '*blaTEM-1* reporter was subcloned as a *Bgl*II- *Bam*HI fragment into each of the multi-copy vectors described below. **(i) *Δ*ss, '*blaTEM-1***. '*blaTEM*-1 was digested from pJM114, end-filled with Klenow and cloned into *Msc*I cut pMV261. The resulting plasmid, pJES102, contains the '*blaTEM-1* reporter without a fused signal sequence cloned downstream

of the *hsp60* promoter. **(ii) ss*plcB-'blaTEM-1***. The *'blaTEM-1* fragment was subcloned into *Bam*HI cut pMB222. The resulting plasmid, pJES101, contains an in-frame fusion of DNA encoding the signal sequence of PlcB/Rv2350c (ss*plcB*) to '*blaTEM-1* under the control of the *hsp60* promoter. **(iii) ss*mpt63-'blaTEM-1***. The *'blaTEM-1* fragment was subcloned into *Bam*HI cut pMB227. The resulting plasmid, pJES103, contains an in-frame fusion of ss*mpt63* (*Rv1926c*) to '*blaTEM-1* under the control of the *hsp60* promoter. **(iv) ss*mpt83-'blaTEM-1***. DNA encoding the signal sequence and the first 31 amino acids of the mature *M. tuberculosis* Mpt83 (Rv2873) protein along with the native *mpt83* promoter (20) was amplified from *M. tuberculosis* genomic DNA using the following primers: mpt83HindIIIF (5'-CAAGCTTCGTCGGATCCGTGGTAGGGGATGTC) and mpt83HindIIIR (5'-CAAGCTTCGGGGTCAGCCATTGCCGCCGTGG) and cloned into the pCR2.1 vector (Invitrogen) to generate pJES125. A *Hin*dIII fragment from pJES125, carrying ss*mpt83* and upstream genomic sequence, was cloned into *Hin*dIII cut pJES128 (Table 3.1). The resulting plasmid, pJES129, contains an in-frame fusion of ss*mpt83* to '*blaTEM-1* under the control of the native *mpt83* promoter (P_{mpt83}).

Protein quantification by immunoblot

Whole cell lysates of *M. tuberculosis* strains were prepared as described previously (2) with the following modifications. *M. tuberculosis* cultures were grown in 5 ml volumes to mid-exponential phase. The cultures were washed twice and resuspended in PBS 0.02% Tween 80. An equal volume of 10% formalin was added to the washed cultures, which were then incubated at room temperature for 1 hour with frequent mixing by inversion. The formalin fixation step was necessary to kill *M. tuberculosis* before further processing. Bacteria were then harvested by centrifugation at 3000 RPM, washed once in PBS 0.02% tween to remove residual formalin, and bead-beaten lysates were then obtained from each sample. Protein concentration for each lysate was measured using a bicinchoninic acid protein quantification kit (Pierce). Lysates were boiled for 10 minutes, subjected to SDS-PAGE and immunoblots were performed using standard conditions. Primary antibodies specific for BlaTEM-1

were used at a concentration of 1:5000 (QED Biosciences), and horseradish peroxidase-conjugated anti-mouse secondary antibodies were used at a concentration of 1:20,000. Bands were visualized using Western Lightning Chemiluminescent Reagent Plus (PerkinElmer) and quantified using ImageJ Image Processing and Analysis software (http://rsb.info.nih.gov/ij/). Whole cells lysates with the highest level of expression were diluted to enable direct comparison of all hybridization signals on a single blot. The comparative quantification was determined by measuring pixel density of an equal area for each blotted lysate in duplicate. Signal intensity per μg of whole cell lysate loaded was determined and is reported as the amount relative to protein detected in the 'BlaTEM-1 expressing strain.

Macrophage infections

THP-1 cells were maintained in RPMI (Gibco)/10% heat inactivated fetal calf serum (FCS) at 37 °C and 5% CO_2. To prepare THP-1 monolayers for infection, cells were spun down at 300 *g*, washed once in RPMI, then resuspended in RPMI /10% FCS at a concentration of 1×10^6 cells/ml. Cells were seeded into 8-well tissue culture slides at 2×10^5 cells/well and treated with phorbol myristate acetate (PMA) at a final concentration of 50 ng/ml for 48 hours.

M. tuberculosis was grown to mid-exponential phase OD_{600} of 0.5-1.0. Immediately prior to infection, the bacterial culture was pelleted, washed once in PBS containing 0.05% Tween 80 (PBS-Tw), and resuspended in an equal volume of PBS-Tw. The culture was then briefly sonicated to break up clumps of bacteria, diluted in RPMI/10%FCS medium and added to the THP-1 monolayer at m.o.i. = 0.1. THP-1 monolayers were infected with *M. tuberculosis* strains for 4 hours at 37 °C and 5% CO_2. Overlaying medium was then removed, the monolayers were washed 3 times with RPMI to remove non-cell associated bacteria, and triplicate wells were lysed and plated to determine uptake (day 0 time-point). The infected monolayers were then overlaid with RPMI/10% FCS, or RPMI/10% FCS supplemented with carbenicillin and maintained at 37 °C and 5% CO_2. At 3 days post infection, the overlying medium was replenished with RPMI/10% FCS media or media supplemented with carbenicillin, as appropriate. On days 1, 3 and 5 post-infection, triplicate wells for each infection were

washed to remove antibiotic and lysed with 0.05% SDS. The resulting lysates were diluted and plated on 7H10 agar to enumerate intracellular bacteria during the course of infection. On day 0 and day 5 of the infection, cell lysates were also plated on 7H10 agar supplemented with 50 μg/ml carbenicillin. This demonstrated that selection of spontaneous β-lactam resistant mutants did not occur during the course of infection. To determine the appropriate carbenicillin concentration necessary to kill intracellular bacteria, THP-1 infection experiments were performed with a range of antibiotic concentrations (Fig. 3.4b). Carbenicillin at 1 mg/ml was determined to be the lowest concentration of antibiotic that caused optimal killing of sensitive intracellular *M. tuberculosis* and was used in subsequent experiments.

RESULTS

'BlaTEM-1 is exported by Sec and Tat signal sequences in *M. tuberculosis*

β-lactamase is an ideal reporter for protein export because it must be localized beyond the bacterial cytoplasmic membrane to effectively protect the bacterium from β-lactam antibiotics. Therefore, it can be used in protein fusions to identify proteins that are extracytoplasmic. An attractive feature of a β-lactamase reporter is that a selection for β-lactam resistant colonies can be performed, as opposed to a more labor-intensive screen. In the past, we showed that the endogenous β-lactamase of *M. tuberculosis* BlaC can function as a reporter of export exclusively by the Tat pathway when expressed in the β-lactam sensitive *ΔblaC* mutant of *M. tuberculosis* or *ΔblaS* mutant of *M. smegmatis* (25). Since the Sec pathway is likely responsible for the majority of protein export in *M. tuberculosis*, we were interested in utilizing a β-lactamase reporter that additionally works with Sec exported proteins. For this reason, we tested the *E. coli* TEM-1 β-lactamase (BlaTEM-1) which has been used in other bacteria to report on proteins exported by Sec, Tat, Type II and Type III secretion systems (5, 7, 37, 39).

A series of multi-copy kanamycin-marked *'blaTEM-1* plasmids were constructed and electroporated into the *ΔblaC* mutant of *M. tuberculosis* (Fig. 3.1). The resulting kanamycin resistant strains were tested for the ability to grow in the presence of 50 μg/ml of the β-lactam carbenicillin.

When the truncated *'blaTEM-1* reporter without a signal sequence was expressed in the *ΔblaC* mutant of *M. tuberculosis,* the strain remained carbenicillin-sensitive. In fact, no colonies of the strain expressing the truncated 'BlaTEM-1 grew on agar containing carbenicillin even after extended incubation (Figs. 3.1 and 3.2). However, expression of a hybrid protein comprised of a Sec signal sequence from Mpt63, a well-established secreted protein of *M. tuberculosis* (17, 24), fused to 'BlaTEM-1 (ssMpt63-'BlaTEM-1) protected the *ΔblaC* mutant from carbenicillin, as was evident by the ability of this strain to grow on carbenicillin agar plates (Figs. 3.1 and 3.2) . We similarly tested a fusion protein in which the Sec signal sequence of a proven cell wall-associated lipoprotein, Mpt83 (16), was fused to 'BlaTEM-1. This construct also conferred β-lactam resistance to *ΔblaC M. tuberculosis* (Fig. 3.1). Of note, the ssMpt83-'BlaTEM-1 fusion protein also included the first 31 amino acids of the mature Mpt83 protein as well as the native *mpt83* promoter which is reported to be active at very low levels *in vitro* (16, 34).
Finally, we tested the signal sequence of PlcB, a proven cell wall-associated phospholipase C, for the ability to promote export of enzymatically active 'BlaTEM-1 (19, 31). PlcB has a predicted Tat signal sequence, and the ssPlcB-'BlaTEM-1 fusion also allowed *ΔblaC M. tuberculosis* to grow in the presence of carbenicillin (Fig. 3.1).

To determine whether the ssPlcB-'BlaTEM-1 fusion was exported by the Tat pathway, it was tested in *ΔblaS M. smegmatis* and in a *ΔtatA ΔblaS M. smegmatis* double mutant (25) in two independent experiments. When the ssPlcB-'BlaTEM-1 fusion protein was expressed in *ΔblaS M. smegmatis*, 92% of colonies were carbenicillin resistant. However, when the same construct was expressed in the *ΔtatA ΔblaS* mutant only an average 7% of colonies were carbenicillin resistant

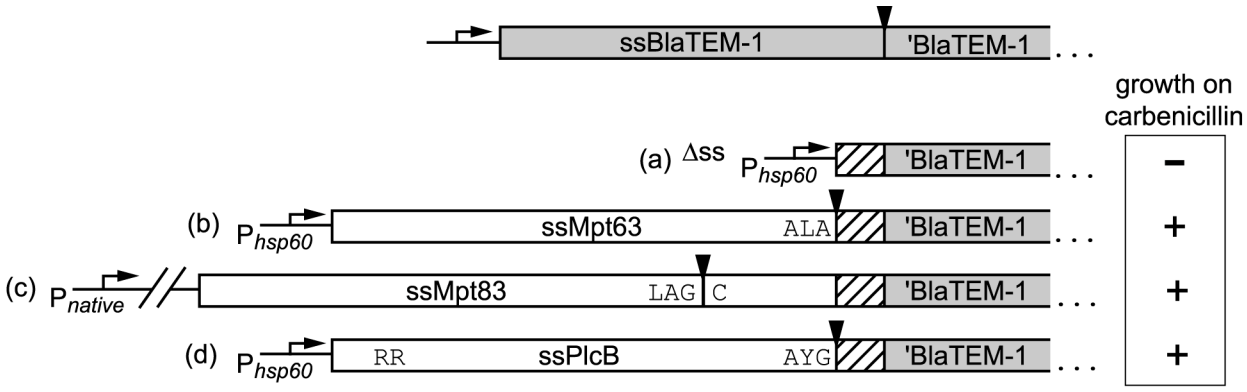

Fig. 3.1. Schematic representation of signal sequence-'BlaTEM-1 fusion constructs. Mycobacterial shuttle plasmids were designed to encode fusion proteins of *M. tuberculosis* peptide sequence (open boxes) with a truncated 'BlaTEM-1 protein (gray boxes) lacking its native signal sequence. The hatched boxes indicate plasmid-derived peptide sequence that is present as a result of the cloning process. The constructs were driven off the constitutive *M. tuberculosis hsp60* promoter for (a) pJES102/*'blaTEM-1*, (b) pJES103/ss*mpt63-'blaTEM-1*, and (d) pJES101/ss*plcB-'blaTEM-1*. The native *M. tuberculosis* promoter located upstream of the *mpt83* operon was used to drive expression of (c) pJES129/ ss*mpt83-'blaTEM-1* (promoters indicated by arrows). Signal peptidase cleavage sites are indicated by arrowheads and by the AxA/G recognition motif for PlcB and Mpt63, and the LAGC lipobox recognition motif for Mpt83. Diagram not to scale; ss, signal sequence.

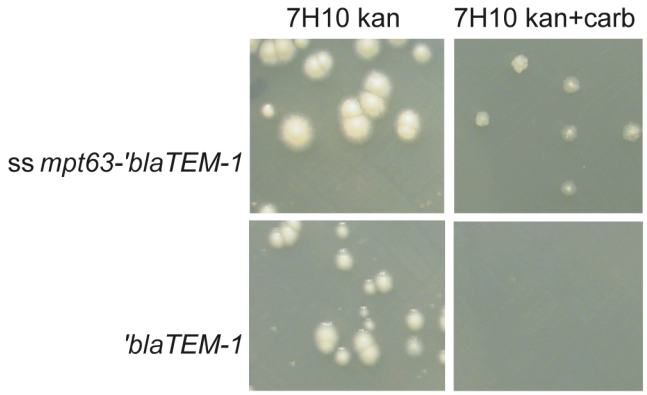

Fig. 3.2. 'BlaTEM-1 does not provide β-lactam-resistance to *ΔblaC M. tuberculosis.* Plasmids encoding the indicated *'blaTEM-1* fusions were electroporated into *M. tuberculosis ΔblaC.* The resulting strains were then plated on 7H10 plates supplemented with either kanamycin and 0.05% tween or kanamycin and carbenicillin without tween. Plates were inspected for growth following 21-25 days of incubation. Not shown are colonies expressing ssMpt83-'BlaTEM-1 and ssPlcB-'BlaTEM-1; growth on plates containing carbenicillin for these strains was similar to that conferred by ssMpt63-'BlaTEM-1.

indicating that the Tat pathway functions in the export of this fusion protein. To show that a functional Tat pathway was not required for export of the Sec signal sequence-'BlaTEM-1 fusion, we similarly evaluated export of ssMpt63-'BlaTEM-1. When expressed in *ΔblaS* and the *ΔtatA ΔblaS* mutants, ssMpt63-'BlaTEM-1 conferred carbenicillin resistance to 90% and 95% of colonies, respectively. This indicated, as expected, no role for the Tat pathway in exporting a Sec signal sequence-'BlaTEM-1 fusion.

In each example where a *M. tuberculosis* signal sequence (Sec or Tat) was fused to 'BlaTEM-1, *ΔblaC M. tuberculosis* was protected from β-lactam attack. To demonstrate that the inability of the 'BlaTEM-1 reporter lacking a signal sequence to protect against carbenicillin was due to lack of export, as opposed to lack of expression, whole cell extracts of 'BlaTEM-1 expression strains were prepared and assayed for cell-associated β-lactamase. To test for enzyme activity, we used the chromogenic β-lactam nitrocefin, which turns red following cleavage by β-lactamase (27). During a 15 minute incubation the nitrocefin was hydrolyzed by all strains expressing 'BlaTEM-1 constructs, while *ΔblaC M. tuberculosis* demonstrated no activity, similar to PBS alone (data not shown). Importantly, β-lactamase activity was detected with the truncated 'BlaTEM-1 reporter lacking a signal sequence. In fact, the lysate from the 'BlaTEM-1 strain converted nitrocefin to the red product almost instantaneously and faster than any other strain tested. We similarly detected β-lactamase activity in whole cell lysates of *ΔblaC M. tuberculosis* expressing the 'BlaC reporter lacking its native signal sequence.

We also compared the level of each 'BlaTEM-1 fusion protein present in whole cell lysates from the respective *M. tuberculosis* strains by immunoblots with antibodies specific for BlaTEM-1. This revealed a wide variation in the amount of 'BlaTEM-1 protein produced by the different strains (Fig. 3.3). The non-exported 'BlaTEM-1 expressed off the *hsp60* promoter (P_{hsp60}) was the most abundant protein detected. P_{hsp60} is considered a relatively strong promoter and is, therefore, present on many mycobacterial shuttle vectors (42). In comparison, the P_{hsp60} driven ssPlcB-'BlaTEM-1 and ssMpt63-'BlaTEM-1 were expressed at lower levels (59% and 0.9% of the level of the non-exported 'BlaTEM-1

construct, respectively). Since *mpt83* is expressed at relatively low levels *in vitro* we expected the ssMpt83-'BlaTEM-1 fusion to be weakly expressed (16, 34, 38). In fact, it was nearly undetectable by immunoblot, present at only 0.4% of the amount of non-exported 'BlaTEM-1 construct. The bands detected on the immunoblot are in general agreement with the predicted molecular weight of the expressed proteins. 'BlaTEM-1, lacking a signal sequence, has a predicted size of 28 kDa. Since whole cell lysates were analyzed in these experiments it is possible to see processed protein and/or uncleaved cytosolic precursor, which may explain the larger sized ssPlcB-'BlaTEM-1 product. The signal sequences of PlcB and Mpt63 would add approximately 3 and 4 kDa, while the Mpt83 signal sequence and fused portion of the mature protein would add approximately 11 kDa, if left intact.

These observations suggested that even though 'BlaTEM-1 does not promote growth in the presence of carbenicillin, a significant amount of β-lactamase was produced and accumulated within the bacterium. Together, our results indicated that in *ΔblaC M. tuberculosis* 'BlaTEM-1 must be exported to confer protection against β-lactam antibiotics, that β-lactam resistance can be used to report on export, and that this reporter can be exported by Sec or Tat signal sequences and is compatible with different levels of expression.

The *ΔblaC* mutant of *M. tuberculosis* is sensitive to β-lactams during intracellular growth in human THP-1 cells

β-lactam antibiotics can be used for clinical treatment of intracellular pathogens such as *Listeria monocytogenes* (33), and have been shown to reduce the population of phagocytosed *Staphylococcus aureus* (1). This indicates that β-lactams can enter macrophages and inhibit intracellular growth of some bacteria. The *ΔblaC* mutant of *M. tuberculosis* is sensitive to β-lactams *in vitro*, and we set out to test if this mutation also makes *M. tuberculosis* susceptible to β-lactams during growth in host cells.

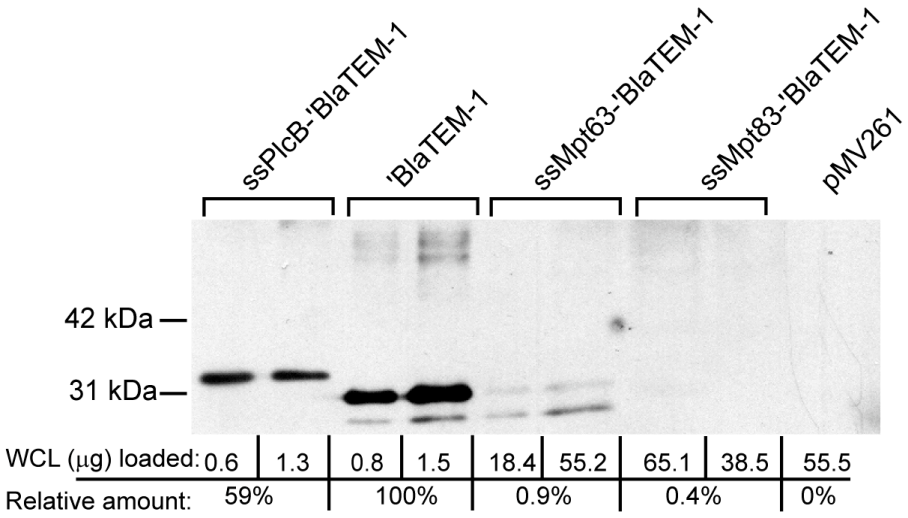

Fig. 3.3. 'BlaTEM-1 fusion proteins are detected at different amounts in *M. tuberculosis* whole cell lysates. Protein present in whole cell lysates (WCL) from each of the indicated *ΔblaC* strains were separated by SDS-PAGE and immunoblotted using primary antibody specific for BlaTEM-1. Comparative signal was quantified by measuring pixel density of an equal area for each blotted lysate in duplicate. Average signal intensity per μg of WCL is reported as the amount relative to protein detected in the 'BlaTEM-1 expressing strain. Due to the different amounts of protein in each strain, it was necessary to load dilutions of the 'BlaTEM-1 and ssPlcB-'BlaTEM-1 expressing lysates so that signal from less abundant protein fusions could be simultaneously detected. There was no detectable signal with the WCL from the *ΔblaC* mutant carrying empty pMV261 plasmid.

Intracellular growth of the *ΔblaC* mutant was not previously evaluated; therefore, we first tested the ability of this mutant to grow within human monocytic THP-1 cells. THP-1 cells were infected at a m.o.i. of 0.1 with either the *ΔblaC* mutant or the virulent parental H37Rv strain. After a four hour period of infection, the THP-1 monolayer was washed to remove non-cell associated bacilli and fresh media was added back. Growth over a five day period was assessed by plating of infected host cell lysates for viable bacilli. The *ΔblaC* mutant showed no difference in intracellular growth when compared to H37Rv (Fig. 3.4a). Of note, we confirmed that *M. tuberculosis* does not grow in the THP-1 culture medium as previously reported (49).

To determine if the *ΔblaC* mutant was sensitive to β-lactams during intracellular growth, THP-1 cells were infected with *ΔblaC M. tuberculosis* and, following the washes to remove extracellular bacilli, media containing different concentrations of carbenicillin was added to the infected monolayers. After five days incubation, the infected monolayers were washed to remove carbenicillin and lysed to plate for viable bacilli. In the absence of carbenicillin, the *ΔblaC* mutant grew in THP-1 cells as previously seen. However, as the concentration of carbenicillin during the intracellular growth period increased, growth of the mutant diminished. At carbenicillin concentrations of $\geq$ 0.8 mg/ml substantial killing of the mutant was observed (Fig. 3.4b). These results indicated that the *ΔblaC* mutant is sensitive to β-lactam antibiotics during intracellular growth, and it suggested that the β-lactamase reporters could be used to study protein export during intracellular growth. Additional experiments showed that a concentration of 1 mg/ml carbenicillin was sufficient to achieve significant killing of the *ΔblaC* mutant of *M. tuberculosis* in THP-1 cells, and this concentration was used in all subsequent experiments.

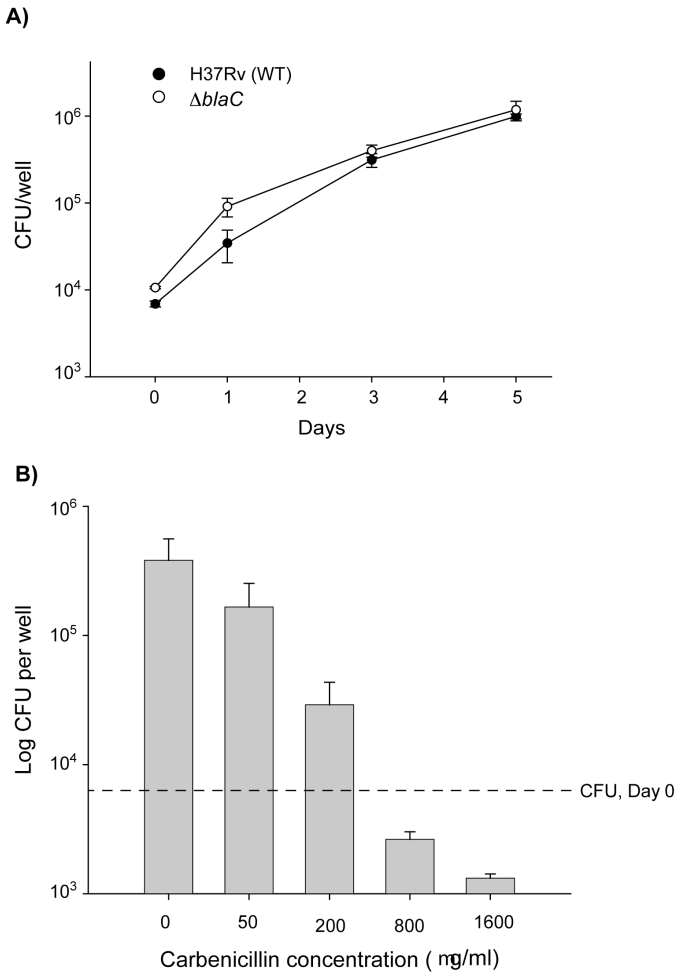

Fig. 3.4. The *ΔblaC* mutant of *M. tuberculosis* does not have a growth defect and is sensitive to β-lactam antibiotic in human THP-1 macrophage-like cells. (a) THP-1 cells were seeded into 8 well chamber slides, and triplicate wells were infected with either WT H37Rv or *ΔblaC M. tuberculosis* at a m.o.i. of 0.1 bacilli per macrophage. At 4 hours (Day 0), 1, 3 and 5 days post infection, infected wells were washed, lysed, and plated for intracellular bacteria. Error bars represent standard error of the mean of c.f.u. in triplicate wells. (b) THP-1 cells were infected with *ΔblaC M. tuberculosis* as in (a). Following a 4-hour uptake period, wells were washed and indicated concentrations of carbenicillin were added to infected wells. Infected cells were lysed and plated at 4 hours (Day 0) and 5 days post infection to enumerate intracellular bacteria. Dashed line represents average intracellular CFU at 4 hours post infection. Error bars represent standard error of the mean of quadruplicate wells combined from two replicates.

Export of β-lactamase protects intracellular *ΔblaC M. tuberculosis* from β-lactam antibiotics

A reporter system that works with intracellularly growing *M. tuberculosis* would be of great value for identifying exported proteins that are expressed and exported only during infection. Having shown that the *ΔblaC* mutant was sensitive to β-lactams during intracellular growth, we tested if β-lactamase could be used to report on protein export by *M. tuberculosis* growing in host cells. We tested fusion proteins expressing the ‘BlaC and ‘BlaTEM-1 reporters for the ability to protect the *ΔblaC* mutant in β-lactam treated THP-1 cells. In each experiment we compared an exported fusion protein to the truncated reporter alone.

To test the ‘BlaC reporter, which works with Tat exported proteins only, THP-1 cells were infected with the *M. tuberculosis ΔblaC* mutant expressing ssPlcB-‘BlaC or ‘BlaC only. Media with or without 1 mg/ml carbenicillin was added and the course of infection was monitored over a five day period. In the absence of carbenicillin, both strains grew in THP-1 cells during the course of the experiment. However, in the presence of carbenicillin, the strain expressing the truncated reporter alone did not grow and was reduced by one log over five days while the strain expressing the exported ssPlcB-‘BlaC fusion protein was protected from carbenicillin and grew normally (Fig. 3.5a).

The ‘BlaTEM-1 fusions were similarly tested. When THP-1 cells were infected with *ΔblaC M. tuberculosis* expressing either the exported ssMpt63-‘BlaTEM-1 or the ‘BlaTEM-1 reporter alone, only the strain expressing ssMpt63-‘BlaTEM-1 fusion grew in THP-1 cells in the presence of carbenicillin. The non-exported ‘BlaTEM-1 strain was sensitive to the β-lactam and was reduced in number by one log (Fig. 3.5b). Similarly, *ΔblaC M. tuberculosis* exporting ssMpt83-‘BlaTEM-1 fusion was able to grow in carbenicillin treated THP-1 cells, while the non-exported ‘BlaTEM-1 construct did not confer resistance to the *ΔblaC* mutant (Fig. 3.5c).

These experiments demonstrated that both the Tat specific ‘BlaC reporter and the more permissive ‘BlaTEM-1 reporter can report on protein export while *M. tuberculosis* is growing in β-

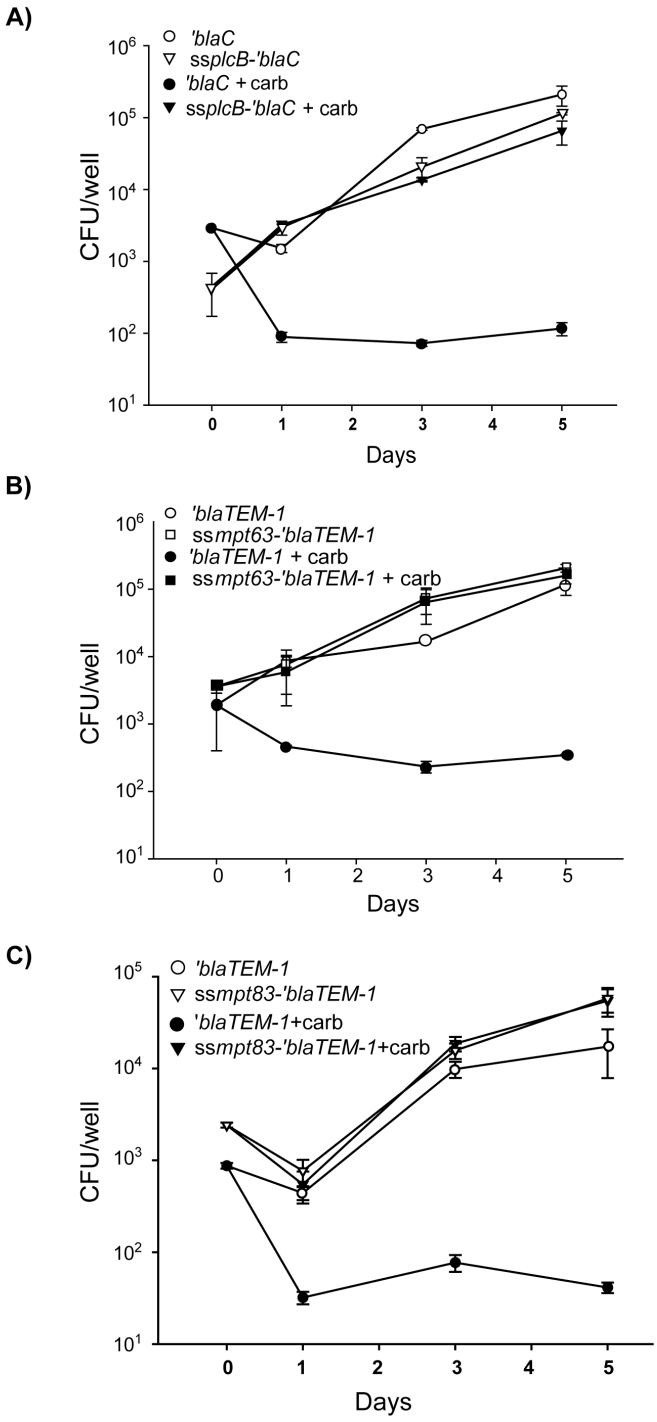

Fig. 3.5. ***M. tuberculosis*** **signal sequences fused to 'BlaC and 'BlaTEM-1 protect intracellular bacilli from β-lactam antibiotics**. THP-1 macrophage like cells were infected in triplicate wells with (**a**) *M. tuberculosis ΔblaC* expressing either ssPlcB-'BlaC or 'BlaC, (**b**) *M. tuberculosis ΔblaC* expressing either 'BlaTEM-1 or ssMpt63'BlaTEM-1 or (**c**) *M. tuberculosis ΔblaC* expressing 'BlaTEM-1or ssMpt83-'BlaTEM-1. Infected cells were then left untreated or treated with 1 mg/ml carbenicillin (carb). Wells were washed, lysed and plated 4 hours (d 0), 1, 3 and 5 days post infection. Each experiment was replicated 3 times in the case of (a) and (b), and 4 times in (c), with similar results.

lactam treated host cells. The use of β-lactamase reporters with intracellular *M. tuberculosis* represents a powerful tool for the study and identification of proteins exported during growth in host cells.

DISCUSSION

The exported proteins of *M. tuberculosis* have been the subject of research attention for some time. This stems from the well-established fact that the majority of bacterial virulence factors and antigens are proteins exported out of the cytoplasm to the bacterial cell envelope or secreted out from the bacterium (13). In fact, there is a growing list of *M. tuberculosis* exported and secreted proteins shown to contribute to virulence or to development of a host immune response (21). Genetic reporters have proven to be powerful tools for identifying these extracytoplasmic proteins. The construction of a β-lactam sensitive *ΔblaC* mutant of *M. tuberculosis* opened the door for using β-lactamases as reporters of protein export directly in *M. tuberculosis*. The 'BlaC reporter can be used as a Tat specific reporter while the 'BlaTEM-1 reporter, shown here, can work with Sec or Tat signal sequences. An advantage of β-lactamase reporters is that they can be used to select for exported fusion proteins, as opposed to more labor intensive screening. In addition, we showed here for the first time that resistance to β-lactam antibiotics can be used to report on protein export during intracellular growth of bacteria. Even in more genetically tractable bacterial pathogens, the identification of proteins exported or secreted from within host cells is a challenge.

Because β-lactams target cell wall modifying enzymes, β-lactamases must be exported in order to protect against these drugs. This export requirement was previously exploited with fusion proteins expressed in *E. coli* and other bacteria grown *in vitro* (5, 22). Here we showed that BlaTEM-1 can also report on protein export directly in *ΔblaC M. tuberculosis*. The three *M. tuberculosis* signal sequences tested in our study are from well-established secreted or cell-wall associated proteins. Mpt63 (Rv1926c, 16kDa protein) has a predicted Sec signal sequence and is one of the four most abundant *M. tuberculosis* proteins secreted into culture media during *in vitro* growth (17). Mpt83 (Rv2873) is a glycosylated

lipoprotein (16, 43) that is exported to the cell wall of *M. tuberculosis*. Mpt83 has a predicted Sec signal sequence with a lipoprotein signal peptidase (LspA) cleavage site and the requisite conserved cysteine for lipid modification. PlcB (Rv2350c, phospholipase C) is a cell wall associated protein of *M. tuberculosis* shown to function in virulence (19, 31). Unlike Mpt63 and Mpt83, PlcB has a predicted Tat signal sequence including a twin-arginine motif (11). Signal sequences from all three of these proteins were able to promote export of a fused 'BlaTEM-1 reporter on the basis of production of β-lactam resistance. Notably, the ssMpt83-'BlaTEM-1 fusion protein was expressed from the native *mpt83* promoter and the fusion protein included the predicted signal sequence plus 31 amino acids of the mature Mpt83 protein. This demonstrated the ability of the reporter to work with different strength promoters and extended protein sequences. It is important to note that even though variable levels of fusion protein were detected in *M. tuberculosis* whole cell lysates as determined by immunoblot, each exported fusion provided sufficient protection against 50 μg/ml carbenicillin while the most abundant 'BlaTEM-1 without an export signal did not confer β-lactam resistance.

Previously, we showed that the PlcB signal sequence is able to drive export of functional 'BlaC in a Tat- and twin RR-dependent manner (25). In *E. coli* the 'BlaTEM-1 reporter works with both Sec and Tat signal sequences (5, 39). The Sec and Tat pathways appear essential in *M. tuberculosis* (2, 35, 36). Therefore, to investigate the mode of export of the ssPlcB-'BlaTEM-1 fusion protein it was tested in *M. smegmatis ΔblaS* and in a *M. smegmatis ΔtatA ΔblaS* double mutant. A 93% reduction in β-lactam resistant colonies was observed in the *M. smegmatis ΔtatA ΔblaS* double mutant. Thus, the Tat pathway is involved in the export of ssPlcB-'BlaTEM-1, although other export pathways participate as well. The signal sequence of PlcB may be promiscuous in targeting the Tat or Sec pathway for export depending on the folded or unfolded nature of a fused reporter element. Similar results were recently shown for some predicted Tat signal sequences in *E. coli* (45).

In addition to working with the Sec and Tat pathways, the'BlaTEM-1 reporter has been used with type II and type III secretion systems of Gram-negative bacteria (7, 37). Since substrates of the

type III secretion system lack conventional N-terminal signal sequences, it remains possible that the ‘BlaTEM-1 reporter will also work with non-conventional exported proteins of *M. tuberculosis*.

An interesting category of exported proteins that has been largely overlooked are those proteins only expressed and/or exported during the course of infection. We hypothesize that these are proteins exclusively exported in the host environment including virulence factors and protective antigens. Further, only a small number of the exported *M. tuberculosis* proteins identified *in vitro* have ever been directly investigated during intracellular growth in host cells (21). For most of these studies, immunomicroscopy was used to localize the proteins in *M. tuberculosis* infected macrophages, which required development of suitable antibodies. We reasoned that if β-lactam antibiotics can reach intracellular *ΔblaC M. tuberculosis,* β-lactamase reporters should additionally work during intracellular growth. β-lactam antibiotics do not normally accumulate in eukaryotic cells; however, antibiotics of this class freely diffuse in and out of host cells (44), and β-lactam antibiotics are used to treat some intracellular bacterial infections (33). More specifically, β-lactams reach intracellular *Staphylococcus aureus* and *Listeria monocytogenes* and prevent growth of these organisms in THP-1 cells (1, 6). Here we showed that *ΔblaC M. tuberculosis* in THP-1 cells was also susceptible to carbenicillin. Thus, BlaC is responsible for *M. tuberculosis* resistance to β-lactam antibiotics during intracellular growth, indicating that the chromosomal *blaC* is a key factor preventing the use of β-lactams to treat *M. tuberculosis* infection.

When the set of exported β-lactamase fusion proteins was tested for the ability to protect *ΔblaC M. tuberculosis* from β-lactam treatment during intracellular growth, all exported fusions conferred resistance. In contrast, the truncated non-exported β-lactamase reporters were not protective. These experiments demonstrated the effectiveness of both ‘BlaC and ‘BlaTEM-1 reporters to identify *M. tuberculosis* sequences that drive export of each reporter during growth within host cells. Because the ssMpt83-‘BlaTEM-1 fusion was expressed from the native promoter, our results indicate that Mpt83, a

protein of unknown function, is expressed and exported during intracellular infection. This result is consistent with the reported induction of *mpt83* in macrophages (38) .

Several approaches have described proteins exported by *M. tuberculosis in vitro*, but a different suite of proteins may be exported during infection of the host. The intracellular β-lactamase reporter system we describe represents a new genetic tool for studying protein export in *M. tuberculosis*. It can be used to directly test the intracellular export of a protein of interest. We also hope to use it in combination with multiple rounds of infection and selection of β-lactam resistant clones from a *M. tuberculosis* fusion library. This should serve to identify the most interesting category of proteins; namely, those that are exported during intracellular growth and missed by alternative methods.

ACKNOWLEDGEMENTS

This research was funded in part by grants from the National Institute of Allergy and Infectious Disease of the NIH (AI54540 and AI070928 to M.B., and AI47311 to M.S.P.) and a developmental grant from the University of North Carolina at Chapel Hill Center for AIDS Research (NIH #9P30 AI 50410-04). J.R.M was supported by an NIH Cell and Molecular Biology training grant (NIH 5-T32-GM008581). J.A.M. was also supported by the Training in Sexually Transmitted Diseases and AIDS NIH NIAID training grant 5T32 AI07001-28 and a Society of Fellows dissertation completion fellowship from the Graduate School at the University of North Carolina at Chapel Hill. We thank Nathan Rigel for assistance with quantitative immunoblots and the members of the Braunstein laboratory for review of the manuscript.

ATTRIBUTIONS

The work described here was a collaborative effort. JRM and MB conceived of the experiments. JRM constructed the reporter system and performed the macrophage experiments. JAM assisted with nitrocefin assays and conducted foundation experiments included in data not shown. MSP constructed the *blaC* mutant of *M. tuberculosis*. JRM, JAM and MB wrote the manuscript. This work has been previously published (*Microbiology*, 2007 Oct; **153**:3350-9), and permission to reprint this work has been granted by the publisher.

REFERENCES

1. **Barcia-Macay, M., C. Seral, M. P. Mingeot-Leclercq, P. M. Tulkens, and F. Van Bambeke.** 2006. Pharmacodynamic evaluation of the intracellular activities of antibiotics against *Staphylococcus aureus* in a model of THP-1 macrophages. Antimicrob Agents Chemother **50:**841-51.

2. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-90.

3. **Braunstein, M., B. J. Espinosa, J. Chan, J. T. Belisle, and W. R. Jacobs, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium* tuberculosis. Mol Microbiol **48:**453-64.

4. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552'phoA in vitro transposition system. J Bacteriol **182:**2732-40.

5. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. Mol Microbiol **4:**1637-44.

6. **Carryn, S., F. Van Bambeke, M. P. Mingeot-Leclercq, and P. M. Tulkens.** 2003. Activity of beta-lactams (ampicillin, meropenem), gentamicin, azithromycin and moxifloxacin against intracellular *Listeria monocytogenes* in a 24 h THP-1 human macrophage model. J Antimicrob Chemother **51:**1051-2.

7. **Charpentier, X., and E. Oswald.** 2004. Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. J Bacteriol **186:**5486-95.

8. **Chubb, A. J., Z. L. Woodman, F. M. da Silva Tatley, H. J. Hoffmann, R. R. Scholle, and M. R. Ehlers.** 1998. Identification of *Mycobacterium tuberculosis* signal sequences that direct the export of a leaderless beta-lactamase gene product in Escherichia coli. Microbiology **144:**1619-29.

9. **Datta, N., and P. Kontomichalou.** 1965. Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae. Nature **208:**239-41.

10. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-20.

11. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-83.

12. **Downing, K. J., R. A. McAdam, and V. Mizrahi.** 1999. *Staphylococcus aureus* nuclease is a useful secretion reporter for mycobacteria. Gene **239:**293-299.

13. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev **61:**136-69.

14. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. Microbiology **151:**521-32.

15. **Guinn, K. M., M. J. Hickey, S. K. Mathur, K. L. Zakel, J. E. Grotzke, D. M. Lewinsohn, S. Smith, and D. R. Sherman.** 2004. Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis.* Mol Microbiol **51:**359-70.

16. **Hewinson, R. G., S. L. Michell, W. P. Russell, R. A. McAdam, and W. J. Jacobs.** 1996. Molecular characterization of MPT83: a seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. Scand J Immunol **43:**490-9.

17. **Horwitz, M. A., B. W. Lee, B. J. Dillon, and G. Harth.** 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **92:**1530-4.

18. **Hsu, T., S. M. Hingley-Wilson, B. Chen, M. Chen, A. Z. Dai, P. M. Morin, C. B. Marks, J. Padiyar, C. Goulding, M. Gingery, D. Eisenberg, R. G. Russell, S. C. Derrick, F. M. Collins, S. L. Morris, C. H. King, and W. R. Jacobs, Jr.** 2003. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. Proc Natl Acad Sci U S A **100:**12420-5.

19. **Johansen, K. A., R. E. Gill, and M. L. Vasil.** 1996. Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. Infect Immun **64:**3259-66.

20. **Juarez, M. D., A. Torres, and C. Espitia.** 2001. Characterization of the *Mycobacterium tuberculosis* region containing the *mpt83* and *mpt70* genes. FEMS Microbiol Lett **203:**95-102.

21. **Kurtz, S., and M. Braunstein.** 2005. Protein secretion and export in *Mycobacterium tuberculosis*, p. 71-138. *In* T. Parish (ed.), *Mycobacterium* molecular biology. Horizon bioscience, Norfolk, UK.

22. **Lee, H. J., and K. T. Hughes.** 2006. Posttranscriptional control of the Salmonella enterica flagellar hook protein FlgE. J Bacteriol **188:**3308-16.

23. **Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi.** 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. J Bacteriol **177:**59-65.

24. **Manca, C., K. Lyashchenko, H. G. Wiker, D. Usai, R. Colangeli, and M. L. Gennaro.** 1997. Molecular cloning, purification, and serological characterization of MPT63, a novel antigen secreted by *Mycobacterium tuberculosis*. Infect Immun **65:**16-23.

25. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-79.

26. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. Trends Microbiol **9:**494-500.

27. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. Antimicrob Agents Chemother **1:**283-8.

28. **Owens, M. U., W. E. Swords, M. G. Schmidt, C. H. King, and F. D. Quinn.** 2002. Cloning, expression, and functional characterization of the *Mycobacterium tuberculosis secA* gene. FEMS Microbiol Lett **211:**133-41.

29. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-40.

30. **Pym, A. S., P. Brodin, L. Majlessi, R. Brosch, C. Demangel, A. Williams, K. E. Griffiths, G. Marchal, C. Leclerc, and S. T. Cole.** 2003. Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. Nat Med **9:**533-9.

31. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-17.

32. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? Nat Rev Microbiol **5:**39-47.

33. **Safdar, A., and D. Armstrong.** 2003. Antimicrobial activities against 84 *Listeria monocytogenes* isolates from patients with systemic listeriosis at a comprehensive cancer center (1955-1997). J Clin Microbiol **41:**483-5.

34. **Said-Salim, B., S. Mostowy, A. S. Kristof, and M. A. Behr.** 2006. Mutations in *Mycobacterium tuberculosis* Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. Mol Microbiol **62:**1251-63.

35. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-79.

36. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

37. **Sauvonnet, N., and A. P. Pugsley.** 1996. Identification of two regions of *Klebsiella oxytoca* pullulanase that together are capable of promoting beta-lactamase secretion by the general secretory pathway. Mol Microbiol **22:**1-7.

38. **Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan, and G. K. Schoolnik.** 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. J Exp Med **198:**693-704.

39. **Stanley, N. R., F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, and B. C. Berks.** 2002. Behaviour of topological marker proteins targeted to the Tat protein transport pathway. Mol Microbiol **43:**1005-21.

40. **Stanley, S. A., S. Raghavan, W. W. Hwang, and J. S. Cox.** 2003. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. Proc Natl Acad Sci U S A **100:**13001-6.

41. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-60.

42. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, and et al.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-60.

43. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-59.

44. **Tulkens, P. M.** 1991. Intracellular distribution and activity of antibiotics. Eur J Clin Microbiol Infect Dis **10:**100-6.

45. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. J Biol Chem.

46. **Wiker, H. G., M. A. Wilson, and G. K. Schoolnik.** 2000. Extracytoplasmic proteins of *Mycobacterium tuberculosis* - mature secreted proteins often start with aspartic acid and proline. Microbiology **146 (Pt 7):**1525-33.

47. **WorldHealthOrganization** 2005, posting date. WHO Information tuberculosis fact sheet. [Online.] http://www.who.int/mediacentre/factsheets/fs104/en/

48. **Zahrt, T. C., and V. Deretic.** 2002. Reactive nitrogen and oxygen intermediates and bacterial defenses: unusual adaptations in *Mycobacterium tuberculosis*. Antioxid Redox Signal **4:**141-59.

49. **Zhang, M., J. Gong, Y. Lin, and P. F. Barnes.** 1998. Growth of virulent and avirulent *Mycobacterium tuberculosis* strains in human macrophages. Infect Immun **66:**794-9.

CHAPTER 4

Development of a reporter transposon to identify exported virulence factors in *Mycobacterium tuberculosis*

Jessica R. McCann, Justin A. McDonough, and Miriam Braunstein

Exported proteins are defined as those exposed on the bacterial cell surface or secreted into the environment. Proteins exported by *Mycobacterium tuberculosis* (*Mtb*) are ideally situated to interact and interfere with host factors. Our objective was to comprehensively identify *Mtb* exported proteins using a β-lactamase reporter for export, and then determine which of these proteins is important for virulence. β-lactamase is an enzyme that cleaves β-lactam antibiotics, and it must be exported beyond the cytoplasm to confer antibiotic resistance. We cloned the β-lactamase BlaTEM-1 lacking its native export signal sequence ('BlaTEM-1) into a transposon, which was then delivered into β-lactam-sensitive Δ*blaC M. tuberculosis*. Only those cells in which *'blaTEM-1* is inserted in-frame with a gene encoding an exported protein will produce an exported β-lactamase fusion leading to β-lactam resistance on agar plates. Because the transposon inserts in the ORF, this strategy was used to identify exported proteins of *Mtb* and simultaneously generate a collection of mutants lacking individual exported proteins. We tested each of the 118 unique mutants for growth defects in murine bone marrow derived macrophages.

We identified eight mutants with macrophage growth defects; four of these mutants are not previously shown to have defects in macrophages. Two mutants with intracellular growth defects were also tested in mice. Our identification strategy provided experimental evidence of export for several uncategorized proteins in *M. tuberculosis*. Further, our approach succeeded in identifying mutants with defects in macrophage growth. Further analysis of the exported proteins important for intracellular growth will lead to a better understanding of *M. tuberculosis* pathogenesis.

INTRODUCTION

Mycobacterium tuberculosis, the causative agent of tuberculosis (TB) disease, is a successful intracellular pathogen that causes nearly 2 million deaths each year (107). To cause disease, *M. tuberculosis* must survive and replicate in macrophages in the lung within phagosomal compartments. One of its survival strategies includes the prevention of normal phagosome acidification following uptake by macrophages (85, 106). A common theme for bacterial pathogens is that factors required for virulence are often exported proteins (39). This is not surprising, as exported proteins are ideally localized to interact and interfere with host factors that would normally protect against infection. We are interested in defining the suite of exported proteins of *M. tuberculosis*, including those found on the surface of the bacterium or secreted into the extracellular environment. In addition to playing important roles in pathogenesis, *M. tuberculosis* exported proteins are also commonly immunogenic. Therefore, they are also potentially protective and/or diagnostic antigens (5, 47). Given their potential for roles in virulence and immune responses, there has been a great deal of interest in proteins exported by *M.*

tuberculosis. However, many of the exported proteins of *M. tuberculosis* are still not experimentally identified as such. Further, even fewer exported proteins have been directly tested for roles in *M. tuberculosis* pathogenesis. Finally, even for those exported proteins with proven roles in virulence not one has an experimentally defined function.

Our first objective was to perform large-scale identification of exported *M. tuberculosis* proteins, and we planned to do this using a transposon that included a reporter for protein export. A reporter enzyme fused to a protein of interest can inform on the subcellular location of that protein by following the activity of the reporter. The use of export reporters is a classic approach to identify proteins that are located beyond the cytoplasmic membrane as well as to topologically map the extracellular portions of proteins with multiple transmembrane (TM) domains. One export reporter that has been used extensively in bacteria is alkaline phosphatase, or PhoA (16, 58, 61, 99). This enzyme is typically exported by the Sec pathway and it is enzymatically active only after transport beyond the cytoplasmic membrane. Substrates of PhoA can be incorporated into agar plates and hydrolyzed by the active enzyme, producing a blue colony for those fusions that promote export of PhoA. Thus, fusions to PhoA can be used in a phenotypic screen to report on protein export beyond the cytoplasmic membrane. Transposon mutagenesis with a PhoA export reporter incorporated into the transposon (TnPhoA) has been successful in identifying exported virulence factors in several bacterial pathogens (97). One striking example is its use in *Vibrio cholerae*, the causative agent of cholera. In this study, *V. cholerae* was mutagenized with TnPhoA and the resulting mutants were screened for blue colonies on plates impregnanted with the PhoA substrate. The blue mutants have transposon insertions in ORFs encoding exported proteins. These mutants

were then tested for their ability to cause disease in suckling mice. In this manner, components of the cholera toxin co-regulated pilus, a major *V. cholerae* exported virulence factor, was discovered (97). However, because *M. tuberculosis* has endogenous enzymes that cleave the PhoA substrate, PhoA cannot be used as a reporter for protein export directly in this pathogen.

β-lactamase is another reporter for protein export that has been used previously in bacteria (17, 24, 75, 76, 94). β-lactamase is an exported enzyme that protects the bacterial cell wall from attack by β-lactam antibiotics, and it must be exported beyond the cytoplasm to be effective (75). The β-lactamase enzyme requires an N-terminal signal sequence for export. When this signal sequence is removed, the enzyme remains cytoplasmic and does not protect against β-lactams (65, 75) (Fig. 4.1). Because *M. tuberculosis* is naturally resistant to β-lactams due to an endogenous β-lactamase (BlaC), β-lactamase reporters were not previously applicable to this pathogen. However, with the recent construction of a β-lactam sensitive *blaC* mutant (40, 41) it became possible to employ β-lactamase reporters directly in *M. tuberculosis*. We previously demonstrated that a non-native β-lactamase, BlaTEM-1 (29), works as an export reporter directly in Δ*blaC M. tuberculosis* (65). Signal sequences derived from *M. tuberculosis* Tat- or Sec-exported proteins are both able to export a 'BlaTEM-1 reporter lacking its own export signals.

We cloned a truncated β-lactamase ('*blaTEM-1*) lacking its own signal sequence into the *mariner*-based transposon *Himar-1* (8, 56, 82), and inserted the reporter transposon into a non-essential region of the mycobacteriophage phAE159. While genetic tools for the modification of the *M. tuberculosis* genome exist, they have been

slow to develop. The organism is difficult to manipulate and not very competent for taking up foreign DNA. One tool at our disposal is the conditionally replicating shuttle phasmid phAE159 (13, 56). Genetic constructs can be cloned into a non-essential region of the mycobacteriophage, which then efficiently delivers foreign DNA to *M. tuberculosis*.

We delivered the transposon via mycobacteriophage infection of Δ*blaC M. tuberculosis*. Transposition that placed '*blaTEM-1* in frame with a gene encoding an exported *M. tuberculosis* protein resulted in production of an exported fusion protein and resistance to β-lactam antibiotics. This strategy allowed us to identify exported proteins directly in virulent *M. tuberculosis* using a positive selection, not a screen. Only those CFU expressing an exported 'BlaTEM-1 fusion protein should be resistant to β-lactams on plates. Because the transposon interrupts the ORF, nearly all the transposon mutants collected will lack the disrupted gene product. Thus, we were able to combine a positive selection for the identification of *M. tuberculosis* exported proteins with collection of *M. tuberculosis* mutants lacking these individual exported proteins.

We planned to comprehensively identify the exported proteins of *M. tuberculosis* using our reporter transposon in a saturating mutagenesis strategy, and then determine which of these exported proteins is required for intracellular growth. Saturating transposon mutagenesis is a powerful tool for making libraries of insertion mutants that can then be tested for phenotypes under different conditions. Two such mutant libraries have previously been made in *M. tuberculosis* (55, 77, 89). The first transposon library

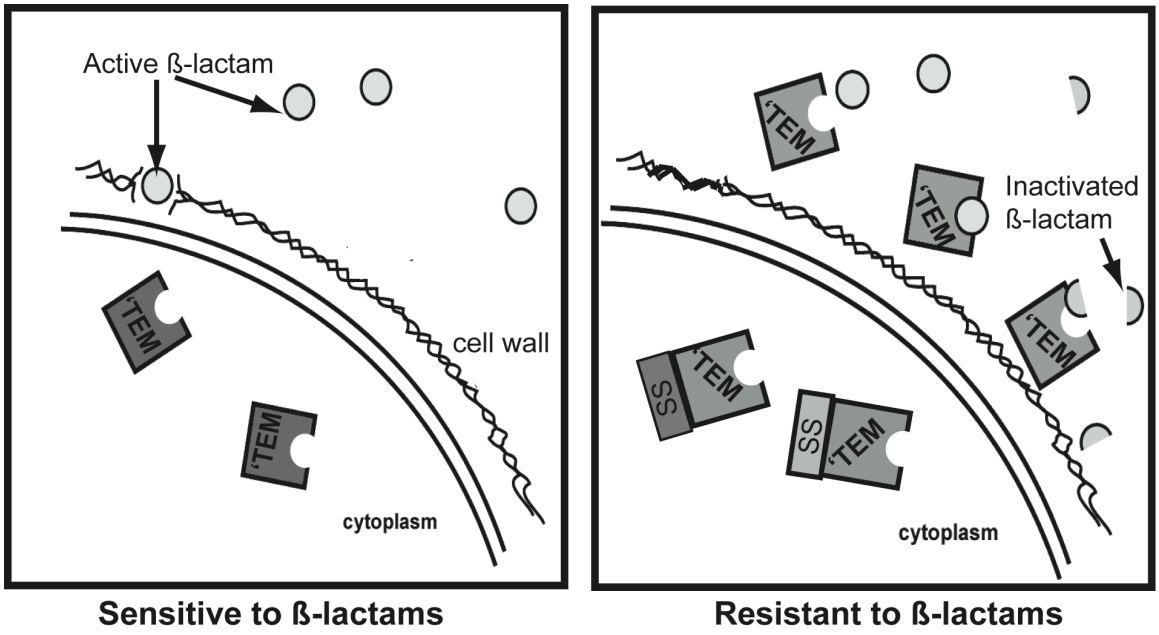

Figure 4.1. 'BlaTEM-1 must be exported beyond the cytoplasmic membrane to confer resistance to β-lactam antibiotics. β-lactam antibiotics (circles) attack the growing cell wall of bacteria. BlaTEM-1 lacking a signal sequence is trapped inside the bacterial cell, and cannot disarm the β-lactam. Fusion of an export signal sequence to '*blaTEM-1* allows for export and results in resistance to β-lactam antibiotics.

was subjected to a technique called Transposon Site Hybridization (TraSH) to predict genes necessary for growth in macrophages or mice. In this technique, mice or or macrophages are infected with libraries of transposon mutants. Following infection, the strains present in the input pool are compared to strains present following infection based on microarray hybridization. Mutants missing or underrepresented in the output pool are predicted to be attenuated (77, 89). A second library was made in a similar manner, but selectively pre-screened for transposon insertions likely to interrupt function of the entire gene in question. This strategy, called Designer Arrays for Defined Mutant Analysis, or DeADMAn, also evaluated pools of transposon mutants simultaneously for those attenuated in mice (55).

A third transposon mutagenesis study conducted in the vaccine strain *M. bovis* bacille Calmette Guerin (BCG) complements the two saturating mutagenesis studies described above. Pools of *M. bovis* BCG transposon mutants were evaluated for their ability to replicate in macrophages following three rounds of infection (92). As in the TraSH and DeADMAn assays, mutants underrepresented in the outpoot pool by microarray hybdridazation are considered to be attenuated (92).

When large pools of mutants are simultaneously evaluated, it is imperative that predictions be validated with careful follow-up experiments. In addition, the use of mutant pools may mask phenotypes of attenuated strains due to trans-complementation by neighboring bacteria. This is a particular concern for mutants that lack secreted virulence factors. In addition, some phenotypes may be exaggerated due to competition between strains during co-infections. Because exported proteins are good candidates for virulence factors, we hypothesized our library of mutants would be enriched for strains

with intracellular growth defects in macrophages. Our second objective was to screen our library of mutants in macrophages to identify exported proteins that are required for *M. tuberculosis* to grow inside cells. Each of our β-lactam resistant transposon mutants was screened for growth in murine derived bone marrow macrophages.

Screening of our transposon mutant library complements these other genome-wide strategies that test pools of mutants. By screening mutants individually, there is more confidence in the experimental outcome. Out of the 118 unique ORFS able to export β-lactamase fusions in *M. tuberculosis*, we identified 8 exported proteins that were required for growth in macrophages. Four of these had not previously been implicated as important for intracellular growth in the approaches described above or any prior studies. Thus, our strategy was successful in identifying new exported virulence factors in *M. tuberculosis*. In addition, only 25 out of 118 exported fusions we identified are in ORFs with a predicted function. For many of the proteins we identified, their ability to promote export of 'BlaTEM-1 is the first experimental evidence of their export. Such proof of export will help to describe the function of many of these proteins.

METHODS

Bacterial strains and culture methods

All bacterial strains and plasmids are listed in Table 4.1. *M. tuberculosis* strains were cultured in Middlebrook 7H9 or 7H10 supplemented with 0.5% glycerol, 0.05% tween and 1x ADS (0.5% bovine serum albumin, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl). When applicable, antibiotics were used at the following concentrations: 50 μg/ml hygromycin (Roche), 20 μg/ml kanamycin (Acros Chemicals), and carbenicillin (Sigma) was used at 20 μg/ml for selection of library mutants or 50 μg/ml carbenicillin

for maintentance of control plasmids. *E. coli* strains were grown in Luria-Bertani medium (LB, Fisher) supplemented with antibiotics when necessary as follows: 150 μg/ml hygromycin, 40 μg/ml kanamycin, and 100 μg/ml carbenicillin.

Construction of Tn'blaTEM-1

Standard molecular biology techniques were used (87). All plasmids are listed on Table 4.2. Unless otherwise stated, all PCR reactions were done with the Expand Hi-fidelity PCR kit (Roche) and 2-5% DMSO was included in all PCR reactions. All PCR products were verified by sequencing, and sequencing was done by Eton Bioscience, Inc. (San Diego, CA, and Research Triangle Park, NC) or the UNC-CH Automated DNA Sequencing Facility (Chapel Hill, NC). The transposon Tn'blaTEM-1 was constructed as follows. The *blaTEM-1* gene was PCR amplified without its signal sequence ('*blaTEM-1*) from pJES102 (65) using primers XbaTem-1F (5'-ATCTAGACCAGAAACGCCTGGTGAA-3') and Spe1TemRev (5'-GACTAGTGCTGGATCCGCAATTGTCTTGG-3'). The resulting product was Spe1TemRev (5'-cloned into pCR2.1 (Invitrogen), creating pJES122. pJES122 was digested with *Spe*I and *Xba*I to liberate a '*blaTEM-1* fragment. The plasmid pSH1 (Table 4.2) (56) carries the *Himar-1* transposon, as well as the mariner transposase (Tnpmar) and an *E. coli oriE* origin of replication. Within the inverted repeats of Himar-1 are an *oriR6K* origin of replication and a hygromycin resistance gene. The *oriR6*K origin allows for replication in λpir *E. coli* strains, and it facilitates recovery of genomic regions surrounding the transposon following insertion into the genome. Our objective was to clone the truncated '*blaTEM-1* directly downstream of the left inverted repeat at the

extreme end of the transposon. To achieve this, pSH1 was digested with *Xba*I, resulting in a 2.1 kb fragment and a 3.1 kb fragment. The 2.1 KB fragment was ligated to the '*blaTEM-1*

Table 4.1 Bacterial strains used in this study

Strain	Genotype	Source
H37Rv	*Mycobacterium tuberculosis* virulent lab strain	ATCC
PM638	*M. tuberculosis* H37Rv Δ*blaC*	(41)
MBTB126	PM638::pJES137	This study

Table 4.2 Plasmids used in this study

Plasmid	Genotype	Description	Source
pCR2.1	*bla aph* ColE1	TA cloning vector	Invitrogen
pMV261.*kan* pMV306.kan	*aph* P_{hsp60} *oriM* ColE1	Multicopy mycobacterial shuttle plasmid	(65)
pJES102	*aph* P_{hsp60}-*'blaTEM-1* (*E. coli*) *oriM* ColE1	*'blaTEM-1* from pCR2.1 cloned into pMV261.kan	(65)
pSH1	Tn*Himar-1*, oriR6Kg, HygR	Shuttle vector carrying the Himar-1 transposon	(56)
pJES123	*hygR* oriR6*K, 'blaTEM-*	Intermediate cloning vector	This study
pJES124	*hygR* oriR6K, ColE1*, aph, tn'blaTEM-1, TnpMar*, λcos	Shuttle vector carrying Tn'blaTEM-1	This study
pJES137	*hyg int attP* ColE1 P_{hsp6} '*blaTEM*-1	Integrating hyg marked vector with constitutively expressed'BlaTEM-1	This study

fragment, resulting in pJES123 (Table 4.2). This ligation abolished the *Xba*I site at the 3' end of the '*blaTEM-1* fragment and contained the hygromycin gene. pJES123 was digested with XbaI and ligated to the 3.1 kb fragment of pSH1 containing the transpoase gene. In this manner, '*blaTEM-1* was cloned directly downstream of the left inverted repeat of the transposon in the final vector, pJES124 (Fig. 4.2).

Construction of Tn'blaTEM-1 delivery phage

The plasmid pJES124, carrying Tn'BlaTEM-1, was incorporated into mycobacteriophage phAE159 as previously described, with some modifications (13). phAE159 phasmid was purified using a standard alkaline lysis protocol (87). Resulting

phasmid DNA was digested with *Pac*I, and, following heat inactivation of the enzyme, ligated to *Pac*I linearized pJES124 in a ligation μg ratio of 10:1 phage:vector. The ligated phasmid was then packaged into λ phage heads via the λ cos sites present on pJES124 using the Max Plax kit (Epicentre). The λ packaging step allows recovery of the newly ligated phasmid following transduction into *E. coli*. Phasmid DNA was then isolated from the resulting hygromycin resistant *E. coli*, and digested with *Pac*I to ensure presence of pJES124. Phasmid DNA encoding phAE159::*tn'blaTEM-1* was then electroporated into *M. smegmatis*. Following electroporation, cells were recovered in growth medium for 30 min at 30°C, the permissive temperature for phage replication. Then, aliquots of electroporated cells were mixed with *M. smegmatis* grown to late exponential phase and added to top agar. The top agar mixture was poured onto a solid agar, and incubated at 30°C for two to three days, at which point plaques were visible. From a single plaque, a high titer lysate of phAE159::*tn'blaTEM-1* of at least 10^{10} plaque forming units (PFU) per ml was obtained.

Transposon mutagenesis in *M. tuberculosis*.

To create transposon mutants in *M. tuberculosis*, Δ*blaC M. tuberculosis* was grown to mid exponential phase, washed twice in MP buffer (50 mM Tris-Hcl, pH 7.6, 150 mM NaCl, 10 mM $MgCl_2$ and 2 mM $CaCl_2$) to remove any residual tween detergent from the bacteria. The mycobacteriophage lysate of $1x10^{10}$ was prepared and incubated with 10^9 washed Δ*blaC M. tuberculosis* to achieve a multiplicity of infection (MOI) of 10 at 39°C for 6 hours with hourly inversion to mix. The mixture was then plated on either 7H10 hyg 50 μg/ml plates, to estimate the total number of transposition events, or 7H10

hyg 50 μg/ml carb 20 μg/ml, to select for transposition events that placed 'BlaTEM-1 in the same reading frame as exported proteins. Mock-infected bacteria were also plated to control for spontaneous carbenicillin resistance. Individual carbenicillin CFU resulting from transposition were either patched or grown to mid-exponential phase in 7H9 broth supplemented with hygromycin and plated on 7H10 agar with hygromycin and also with and without carbenicillin to confirm β-lactam resistance.

Genomic DNA purification and recovery of transposon insertions with flanking genomic sequence

CFU determined to be β-lactam resistant following transposition were grown to late-exponential phase in 7H9 hyg broth. Cells from 3 ml of culture were spun down and the supernatant was discarded. The remaining pellets were then lysed by vortexing with 0.6 ml of a 3:1 mixture of chloroform-methanol. The lysate was then mixed by vortexing with 0.6 ml of Tris-equilibrated phenol and then mixed by inversion with 0.9 ml of GTC solution (4 M guanidinium thiocyanate, 0.1 M Tris [pH 7.5], 0.5% sarcosyl, with β-mercaptoethanol added to a final concentration of 1% prior to use). The upper phase was collected after centrifugation, and the genomic DNA was precipitated with an equal volume of isopropanol. To determine the site of insertion for each transposon mutant, approximately 1 μg of genomic DNA was digested with *BssH*II and self-ligated. *E. coli* DH5α λpir was transformed with 1 or 2 μl of each ligation, and plated on LB agar supplemented with hygromycin. Plasmids were then purified from the resulting hygromycin resistant colonies, and *M. tuberculosis* genomic DNA flanking the transposon LIR was sequenced from a primer that hybridized to the complementary

strand of '*blaTEM-1*. Resulting sequences were blasted against the *M. tuberculosis* genome for identification. The predicted amino acid sequences of proteins identified as promoting export were analyzed by bioinformatic algorithms that predict export signal sequences: TMHMM (91), TopPred (25), TMPRED (46), SignalP (9), Psort (69), LipoP (49) and TatP (10). Southern blotting and hybridization were done as previously described (14).

Macrophage Infections

Macrophage infections were completed as described (54). Briefly, bone marrow derived macrophages were obtained from 6-24 week old female BL/6 mice. Approximately 24 hours prior to infection, differentiated macrophages were seeded into 8-well chamber slides at $2x10^5$ cells/well. Indicated *M. tuberculosis* strains were inoculated from freezer stocks grown to mid-exponential phase in 7H9 supplemented with hyg 50 ug/ml. Bacteria were washed once in PBS 0.05% Tween 80, resuspended at $1x10^6$ CFU/ml in DMEM supplemented with 10% FBS, non-essential amino acids and L-glutamine, and added to macrophages in chamber slides at an MOI of 1. Macrophages were allowed a 4-hour uptake period, after which the infected cells were washed with warm DMEM 3 times. A subset of wells were then lysed, diluted and plated on 7H10 agar to determine uptake. Remaining infected wells were lysed, diluted and plated 5 days post infection to determine intracellular bacterial growth.

Animal Infections

Female 6-10 week old C57 BL/6 mice (Charles River) were used for aerosol infection studies. Mice were maintained in hot-washed cages in a BSL-3 animal facility

and given food and water ad libitum. The indicated *M. tuberculosis* strains were grown to mid-exponential phase, washed once and resuspended in phosphate-buffered saline (PBS) with 0.05% Tween 80 (PBS-Tw) to a concentration of 1.2×10^7 CFU/ml. The bacterial suspension was placed into the nebulizer jar of a whole-body exposure aerosol chamber (Mechanical Engineering Workshop, Madison, WI). Mice were exposed for 15 min with a chamber purge time of 20 min. At specific time points, four mice were sacrificed from each group by CO_2 asphyxiation. At days 1 and 14, only lungs were removed and homogenized in PBS-Tw with x 0.1 mg/ml cycloheximide (Sigma). At later time points, lungs, livers and spleens were removed and homogenized as described for lungs only above. The homogenates were plated onto 7H10-ADS-glycerol plates with 0.1 mg/ml cycloheximide to obtain CFU counts.

RESULTS AND DISCUSSION

Construction of Tn'blaTEM-1

We previously established that the TEM-1 β-lactamase reporter ('BlaTEM-1) can be exported when fused to *M. tuberculosis* Sec or Tat signal sequences, and it can be used directly in virulent *M. tuberculosis* (65). Our objective was to combine this robust genetic reporter system, with transposon mutagenesis to comprehensively identify exported proteins of *M. tuberculosis* and collect *M. tuberculosis* mutants lacking these proteins. It was imperative that the transposon we used have little site specificity for insertion, so that it could interrupt ORFs throughout the genome with as little bias as possible. The mariner based *Himar-1* transposon inserts at any TA dinucleotide sequence. At least one TA sequence occurs in all but 16 out of the 3995 ORFS in the *M. tuberculosis* genome (20, 55). *Himar-1* is known to distribute randomly throughout the

M. tuberculosis chromosome and has been used in saturating transposon mutagenesis projects (55, 82, 88).

The *Himar-1* transposon was encoded on pSH1 (Table 4.2). Within the inverted repeats of the transposon, there is a hygromycin resistance gene and an oriR6K origin of replication. This origin allowed us to later easily recover and sequence sites of transposon insertion in the *M. tuberculosis* genome. The transposase on pSH1 is encoded outside of the transposon, which ensures that transposition is stable once the transposon inserts. We cloned the TEM-1 β-lactamase lacking its own signal sequence (‘*blaTEM-1*) directly downstream of the left inverted repeat in *Himar-1*, resulting in plasmid pJES124 (Fig. 4.2). The cloning strategy was designed so that there is an open reading frame running through the inverted repeat and into the truncated β-lactamase reporter.

β-lactamase hydrolyzes the β-lactam ring of β-lactam antibiotics, rendering them ineffective in disrupting the bacterial cell wall. Therefore, in order to be protective against β-lactam antibiotics, β-lactamase must be exported to destroy the drug before it reaches its target (Fig. 4.1). We created our transposon so that only those cells in which *'blaTEM-1* is inserted in-frame with a gene encoding an exported protein will be resistant to β-lactam on agar plates due to the production of an exported *M. tuberculosis* protein-‘BlaTEM-1 fusion. In order to use the β-lactamase reporter, we also needed a β-lactam sensitive strain of *M. tuberculosis*. *M. tuberculosis* is naturally resistant to β-lactam antibiotics due to export of its native β-lactamase *blaC*. For this project, we used a β-lactam sensitive Δ*blaC* mutant of *M. tuberculosis* as our background strain (41).

Electroporation of foreign DNA into *M. tuberculosis* is highly inefficient; the best efficiencies achieved with electroporation are 10^4 CFU per μg of DNA (13). Phage

transduction is a far more efficient way to transfer foreign DNA, such as transposons, to *M. tuberculosis*. During phage infection, every bacterial cell present is likely to receive the phage genomic DNA. We chose the TM4-derived temperature sensitive mycobacteriophage phAE159 as our transposon delivery system (8, 13, 56, 100). The phage replicates and is lytic at the permissive temperature of 30ºC. At 39ºC, it is able to translocate its DNA into mycobacteria but does not replicate (8, 13, 56). This mycobacteriophage also has a large deletion in a non-essential region of its genome and can be engineered to carry and deliver a transposon. We engineered the mycobacteriophage phAE159 to deliver our Tn'BlaTEM-1 transposon. The plasmid pJES124, carrying Tn'BlaTEM-1, was linearized with *Pac*I, and cloned into the *Pac*I site of the phAE159 phasmid. The ligated phasmid was then packaged into λ phage heads via the λ cos sites engineered onto pJES124 (Fig. 4.2). The λ packaging step allows recovery of the newly ligated phasmid following transduction into *E. coli*. We isolated phasmid DNA from the resulting hygromycin resistant *E. coli*, and electroporated it into *M. smegmatis*. Following electroporation, cells were recovered in growth medium for at 30°C, the permissive temperature for phage replication. Then, aliquots of electroporated cells were mixed with *M. smegmatis* grown to late log phase and added to top agar. The top agar mixture was poured onto a solid agar, and incubated at 30 degrees for two to three days, at which point plaques were visible. From a single plaque, a high titer lysate of phAE159::Tn'blaTEM-1 of 10^{10} plaque forming units (PFU) per ml was obtained.

Construction of the Tn'blaTEM-1 mutant library in *M. tuberculosis*

Using an MOI of 10, Δ*blaC M. tuberculosis* was infected with the phage phAE159::Tn'blaTEM-1 at the non-permissive temperature of 39°C. A fraction of each phage infection was plated on 7H10 plates with hygromycin and without carbenicillin, which allowed us to estimate the total number of transposon mutants obtained in each phage infection. The average transposition efficiency was 10^{-6} hygromycin resistant CFU/total input CFU, which is consistent with published transposition efficiencies for a similar phage-delivered *Himar-1* transposon in *M. tuberculosis* (59) (Table 4.3). The remainder of each infection was plated on plates supplemented with both hygromycin and carbenicillin to select for the β-lactam resistant mutants. The efficiency for obtaining β-lactam resistant clones from transposition was much lower, on the order of 10^{-8} carbenicillin resistant CFU/total input CFU. All carbenicillin resistant colonies were rechecked by either plating for single colonies or patching on plates supplemented with hygromycin plus carbenicillin. Following confirmation of carbenicillin resistance, genomic DNA was prepared from each mutant and digested with *BssH*II, an enzyme that does not cut within the transposon sequence. The digested genomic DNA was then self-ligated and transformed into the *E. coli* DH5α λ*pir* strain. This strain background allowed us to take advantage of the OriR6K origin of replication on the transposon and

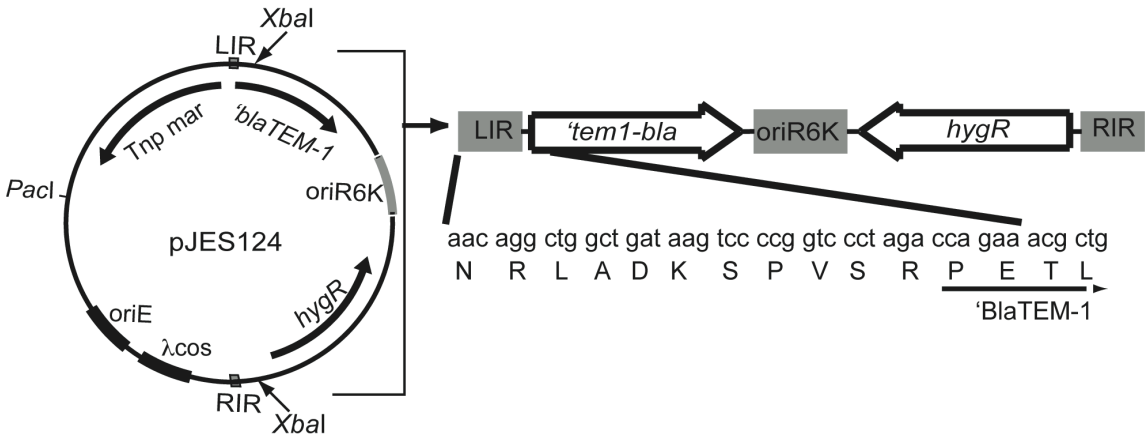

Figure 4.2. Map of transposon Tn'blaTEM-1. Signal sequence-less *'blaTEM-1* was cloned immediately downstream of the inverted repeat (LIR) on a *Himar1*-based mariner transposon on cosmid pSH1. Tnpmar, mariner transposase; hygR: hygromycin resistance gene. The cosmid was cloned into the genome of the temperature-sensitive (Ts) mycobacteriophage phAE159 (56) and packaged into Ts phage. The codon sequence of the inverted repeat is shown in the same reading frame as the starting codons of 'BlaTEM-1.

recover hygromycin resistant plasmids carrying the transposon along with neighboring genomic sequence. A sequencing primer was then used to sequence out from the transposon to identify the site of transposon insertion in the genome. BLAST analysis to the genome of *M. tuberculosis* (1) was used to map the exact location of each transposon insertion (Supplemental Table 4.1, appended at the end of this Chapter).

Our objective was to identify as many exported proteins in *M. tuberculosis* as was possible using the β-lactamase reporter. At the start of the project, we calculated that evaluation of 80,000 transposon insertion mutants would give 95% probability that any given gene of 1 kb in length in the 4.4 Mb *M. tuberculosis* genome would be represented with at least one transposon insertion in the same reading frame as the respective gene. This calculation was based on the following equation: $N = [\ln(1-P)/\ln(1-f)] \times 6$ reading frames, where N = number of transposon insertion mutants required to achieve a probability of P that every ORF of a certain size is represented, and f = fraction of the genome represented by each 1 kb gene (87). After 90 phage infections, 81,000 transposon mutants had been plated onto the carbenicillin containing plates for the selection of transposon insertions in genes encoding exported proteins. From this pool of transposon mutants, 179 colonies were obtained on carbenicillin plates and confirmed to be carbenicillin resistant. Sequence analysis revealed that all 179 of these mutants possessed an in-frame fusion of an ORF to the β-lactamase reporter, which demonstrates the powerful specificity and selection provided by this reporter in *M. tuberculosis*. Some ORFs were identified with the transposon more than once. A total of 118 different ORFs were identified. (Supplemental Table 4.1, appended to the end of this Chapter). As expected from past studies with the *Himar-1* transposon in *M. tuberculosis*, transposon

insertions were obtained throughout the genome (82, 88). Southern blot analysis and evaluation of transposon sequences rescued from each mutant further indicated that the mutants had only a single transposon insertion (Fig. 4.3). Interestingly, in 27 out of 37 cases in which a gene was hit more than once with Tn'blaTEM-1 the insertions occurred at exactly the same site. This occurrence was not due to siblings (amplification of a

Table 4.3 Typical transposition efficiencies[a]

CFU Input/ml	PFU input/ml	Hyg R CFU[b]	Transposition efficiency	Carb R CFU
$1x10^8$	$8x10^9$	150	$1.4x10^{-6}$	4
$8.6x10^8$	$3x10^{10}$	1320	$1.5x10^{-6}$	18
$7.5x10^8$	$2x10^{10}$	470	$6.3x10^{-7}$	2
$3.8x10^8$	$5x10^{10}$	340	$8.9x10^{-7}$	2
$9x10^8$	$5x10^{10}$	1370	$1.5x10^{-6}$	8
Average Transposition Efficiency: $1.2x10^{-6}$			% carb R of total TN insertions[c]: 0.9%	

[a]Sample set of transposition efficiencies chosen from over 90 separate reactions to show range of results

[b]Total number of hygromycin resistant CFU estimated from samples plated for every phage infection

[c]Average percentage of carbenicillin resistant clones calculated from total number of estimated hygromycin resistant transposon mutants.

mutant prior to plating), as nearly all of the examples of replicate insertions were recovered from independent phage infections.

All 118 ORFs identified as able to export active 'BlaTEM-1 fusions possessed predicted export signals, which are the likely explanation for export. Using domain prediction algorithms we classified each ORF as either possessing a cleavable Sec signal sequence, a cleavable lipoprotein signal sequence, a cleavable Tat signal sequence, and/or a TM domain (Table 4.4, see materials and methods for algorithms used). In *M. tuberculosis*, as in many other bacterial pathogens, there also exist unconventional exported proteins that lack obvious signals for export (2, 65, 80). No proteins identified with the reporter transposon fell into this category. Even though all 118 ORFs possessed

predicted export signals, for many of these proteins our identification of them with the β-lactamase reporter transposon is the first experimental evidence they are exported by *M. tuberculosis*. For 80% of the ORFs we identified there is no known or predicted function. For these proteins in particular, their identification with the reporter transposon is significant in establishing a biological property. This will be useful information for

Table 4.4. Description of *M. tuberculosis* Tn'*blaTEM-1* transposon library	
No. total transposon mutants[a]	81,000
No. of confirmed carbenicillin resistant insertions sequenced	179
No. of insertions in ORFs	179
No. of insertions in-frame	179
No. of unique ORFs identified	118
No. of ORFs with Sec signal sequence	43
No. of ORFs with Tat signal sequence	3
No. of ORFs with transmembrane domains	72
No. of ORFs predicted to be lipoproteins	14
No. of ORFs with homologues only in mycobacteria	47
No. of ORFs with homologues only in virulent mycobacteria[b]	17
No. of ORFs with experimentally demonstrated function	8
No. of ORFs reported to function in virulence[c]	28

[a]Total transposon mutants is approximated from plating a fraction of each transposition reaction on agar containing hygromycin only.
[b]Virulent mycobacteria searched: *M. leprae, M. bovis, M. ulcerans,* and *M. avium.*
[c]This includes ORFs identified in genomic approaches for identifying virulence factors (19, 28, 55, 77, 89) and ORFs shown to be required by direct testing of mutants lacking the ORF in macrophages or mice.

determining function of these unknown ORFs. For integral membrane proteins, the site of active β-lactamase insertions provides additional information about protein topology since active insertions must be fused to domains localized on the extracytoplasmic side of the membrane. A striking example of how topology mapping can be aided by our study comes from evaluating the sites of insertions in the MmpL4 protein. MmpL4 is a polytopic membrane protein active that is predicted to have between 11-12 TM domains by bioinformatic algorithms. Experimental evidence is critical to establishing the

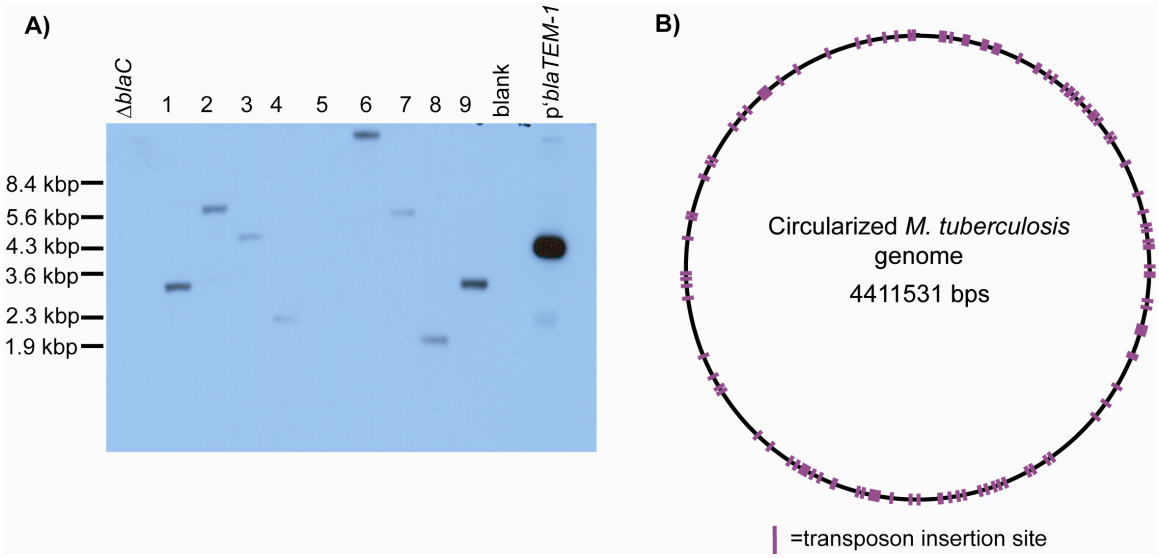

Figure 4.3. Tn'blaTEM-1 inserts randomly throughout the *M. tuberculosis* genome. **A**)Δ*blaC M. tuberculosis* was infected with phAE159:*Tn'blaTEM-1*, and 9 ß-lactam resistant clones (lanes 1-9) were randomly picked and grown to late exponential phase in 7H9 supplemented with ADS, glycerol and 50 μg/ml hygromycin. Genomic DNA was purified from each culture, digested with *Pst*I and separated on a 0.8% agarose gel. Digested DNA was then transferred to a nylon membrane and hybridized to a probe specific to an internal portion of '*blaTEM-1*. **B)** Circularized map of the *M. tuberculosis* genome showing sites of Tn'blaTEM-1 insertions.

extracellular and intracellular domains of an integral membrane protein. Different prediction programs often disagree on the number and location of potential TM domains and protein topology. The MmpL4 protein has a demonstrated role in *M. tuberculosis* virulence and has strong homology to bacterial transporters, but its substrate is unknown (34, 55, 74, 89). We identified three Tn'blaTEM-1 insertions with enzymatic activity in the *mmpL4* ORF. By combining the sites of the active Tn'blaTEM-1 insertions with earlier data obtained using the alkaline phosphatase (PhoA) reporter in *M. smegmatis* to identify exported protein domains of MmpL4 (16), a total of seven extracytoplasmic sites can be delineated in this protein (Fig. 4.4). This experimental data enables development of a relatively refined topology map for this virulence factor of unknown function. Interestingly, it is the two largest soluble domains of MmpL4 that are exposed on the external side of the membrane.

Once we sequenced each transposon insertion in our collection of 180 mutants, we used a Poisson calculation to determine the probability that all exported proteins that could be identified with the β-lactamase reporter transposon had been identified. This number was calculated using the following equation: $f[0] = e^{-m}$, where f[0] represents the still mutable genes that have not yet been recovered, and m is the mean number of hits already recovered per gene (45). This calculation predicted that, following selection from 81,000 transposon mutants, 19% of future carbenicillin resistant CFU collected would still harbor transposon hits in new ORFs not yet identified in our strategy. In fact, in our final infection two out of ten carbenicillin resistant mutants had the transposon inserted in new ORFs not identified earlier. Consequently, continued application of this approach is likely to identify more *M. tuberculosis* exported proteins.

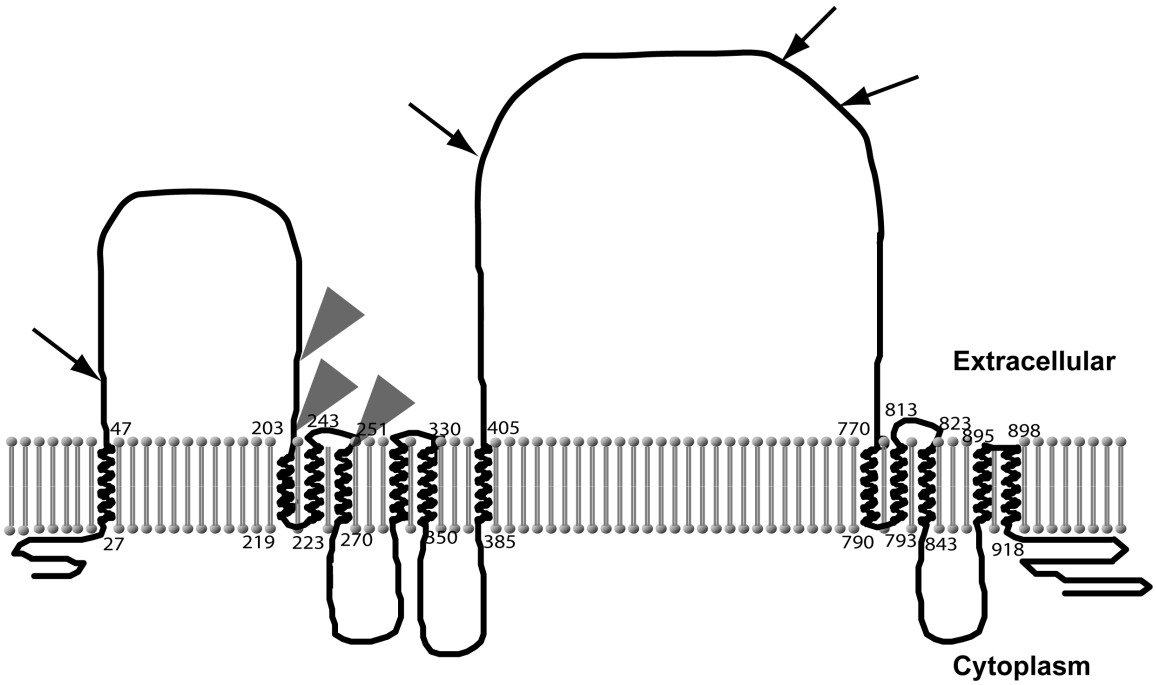

Figure 4.4. Topology map of MmpL4. Transmembrane domains predicted by bioinformatic algorithms were combined with data gathered from Tn'blaTEM-1 insertion sites (triangles) and PhoA fusions [arrows, (16)] to generate a topological map of MmpL4. Amino acid numbers indicate the beginning and end of each predicted transmembrane domain.

Our transposon reporter identification strategy has three significant attributes. The first is that our identification scheme relies on a positive selection – only those mutants with in-frame transposon insertions in ORFs leading to export of the 'BlaTEM-1 reporter will grow in the presence of the β-lactam carbenicillin. The second is that this strategy identifies exported proteins directly in virulent *M. tuberculosis*, and not in a heterologous host expressing *M. tuberculosis* proteins. The third is that the same β-lactam resistant transposon insertions that identify an exported protein most likely disrupt the function encoded by that exported protein. Therefore, at the same time as identifying exported proteins we are creating a library of mutants lacking individual exported proteins that can next be screened for any activity associated with exported proteins.

Transposon insertion mutants with intracellular growth defects

One reason for our interest in identifying exported proteins of *M. tuberculosis* is that these proteins are prime candidates for being factors that interact and interfere with the host. It is often the case that virulence factors in bacterial pathogens are exported proteins (39). Since most transposon insertions will disrupt the function of an ORF, we hypothesized that our insertion mutant library is enriched for mutants lacking exported virulence factors. Our collection of mapped transposon insertions can be used to find mutants with phenotypes, such as defects in virulence, that result from the absence of an exported protein. Upon entry to the lung, *M. tuberculosis* replicates in alveolar macrophages, and this replication is required for the bacterium to cause disease (83). The mechanisms employed by *M. tuberculosis* to survive and grow in the normally hostile macrophage environment are not well understood. By screening our library of mutants

disrupted for exported proteins we proposed that we could identify new genes required for intracellular replication of *M. tuberculosis*.

We screened our transposon insertion mutants for the ability to grow over 5 days in primary murine bone marrow derived macrophages. Each mutant was tested separately in at least two independent experiments. For ORFs that received multiple hits, the transposon mutant with the insertion earliest in the gene was tested as a representative mutant. In each experiment, three independent wells of infected macrophages were assayed for each time point. Macrophage uptake and fold growth over five days for each mutant was compared with that of *M. tuberculosis* MBTB126, a hygromycin resistant strain of Δ*blaC M. tuberculosis* that expresses the truncated 'BlaTEM-1 on an integrating vector. MBTB126 grew at the same rate and to the same level as WT *M. tuberculosis* in macrophages (data not shown). Mutants that were significantly different from MBTB126 in these initial screening assays were tested further. Eight mutants proved to have reproducible and statistically significant intracellular growth defects: *mce1A, mce1B*, *rv0199*, *mmpS5, mce2F, ctaC, ppsB and lppX* (Fig. 4.5). Each of these eight mutants was further evaluated in macrophage infections with added time points (Figs 4.6 - 4.8, 4.10). Of the genes disrupted in these eight insertion mutants, four (*mce1A, mce1B, rv0199*, and *mmpS5*) are previously implicated as having a role in promoting intracellular growth on the basis of TraSH analysis of an *M. tuberculosis* transposon mutant library in macrophages (77). For the other four mutants (*mce2F*, *ctaC*, *ppsB* and *lppX*) there is no previous report of them being defective for intracellular growth in macrophages. Thus, our screening of *M. tuberculosis* transposon mutants in macrophages successfully identified *M. tuberculosis* transposon mutants lacking exported ORFs with growth

defects in macrophages. All eight mutants were tested for their ability to grow in broth culture, and none were defective when compared to MBTB126 (Fig. 4.10 and data not shown).

For these eight exported proteins, only CtaC, PpsB and LppX have a predicted function. For the other five, there is no available amino acid sequence similarity to inform on function. Below, we summarize our data and the current state of knowledge about the disrupted genes and associated genomic locus for each of these eight mutants.

Mce1A, Mce1B, Mce2F

The first Mce protein (Mce1A) was discovered when an *M. tuberculosis* genomic library was transformed into *Escherichia coli* and DNA sequences that promote enhanced entry into non-phagocytic HeLa cells were selected (6). Because of its identification as a protein able to confer cell invasion capability, it was named Mce for *m*ycobacterial *c*ell *e*ntry. When the genome sequence of *M. tuberculosis* was examined, it became apparent that *mce1A* is the third gene in an operon and that four separate *mce* operons (*mce1-4*) exist in the genome. Each operon is composed of a suite of 8-13 similarly arranged genes that are demonstrated to be transcribed as a single polycistronic message (22, 52) (Fig. 6). Near the start of each operon, there are two ORFs with homology to YrbE, a predicted integral membrane protein with similarity to permease proteins of ATP binding cassette (ABC) transporters (21, 53, 98) . In *mce* operons the *yrbE* pair is followed by six genes each encoding members of what is now referred as the Mce protein family (MceA-MceF). The 24 Mce proteins in *M. tuberculosis* share similar domain structures and several conserved motifs.

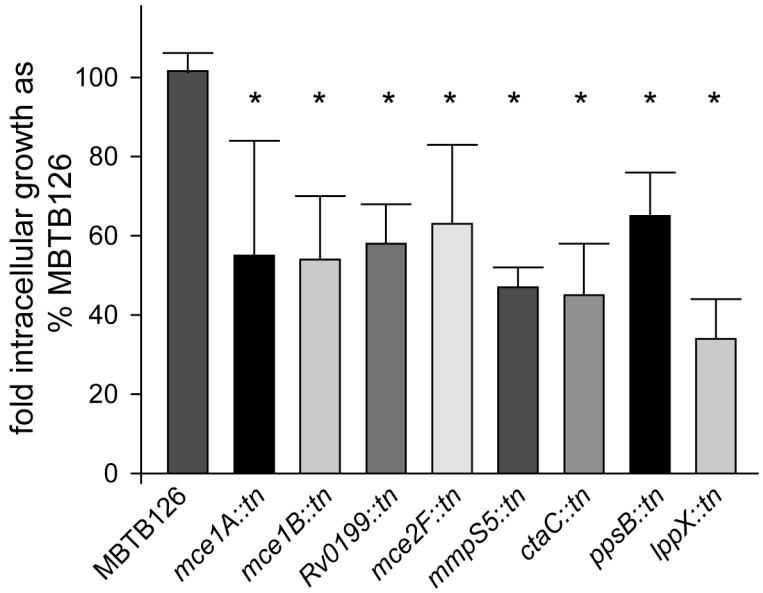

Figure 4.5. Average fold growth of *M. tuberculosis* Tn'blaTEM-1 mutants compared to that of MBTB126. Bone marrow derived macrophages were infected with the indicated strain. At 4 hours (Day 0) and 5 days post infection, infected macrophages from triplicate samples were lysed and plated for intracellular bacilli. The fold growth over five days was calculated for each strain and compared to that of MBTB126 for each respective experiment. The error bars represent standard deviation of the mean for at least three independent experiments per strain. *, $p \leq 0.05$.

Each Mce is predicted to be exported due to the presence of a single N-terminal signal sequence or TM domain (21, 60, 98). Eight of the Mce proteins are predicted to be structurally similar to beta-barrel outer membrane proteins of bacteria (72). The *mce1*, *3* and *4* operons also possess homologous downstream ORFs, which have been referred to as Mce associated proteins (Mas) (21). Using the Tn'blaTEM-1 transposon we identified 14 proteins of *mce* loci as being exported (Fig. 4.6). For nine of these proteins, this represents the first experimental evidence they are exported and expressed by *M. tuberculosis*.

We obtained active transposon insertions in three of the 4 *mce* operons (*mce1, mce2,* and *mce4*). Within the *mce1* operon, we recovered Tn'blaTEM-1 insertions in nearly every gene. Interestingly, only the insertions in *mce1A* and *mce1B* resulted in macrophage growth defects when tested in our assay (Fig. 4.6). However, in the TraSH screen for transposon mutants with intracellular growth defects, insertions in each of the genes spanning from *mce1B to mce1F* lead to underrepresentation in the output pool following macrophage infection (77). In a separate study, an allelic exchange mutant that knocked out expression of the entire *mce1* operon was constructed and directly tested in macrophages. However, this mutant grew better than WT in cultured macrophages and was deemed hypervirulent (90). With the exception of not seeing intracellular phenotypes for mutations in genes downstream of *mce1B*, our results agree with the TraSH analysis of pooled *M. tuberculosis* mutants in macrophages. The disparity between phenotypes in these studies may be due to the nature of insertions in each of the genes and/or how downstream genes are affected in individual mutants. Another possible explanation for the differences in experimental outcomes is that TraSH pooled infections

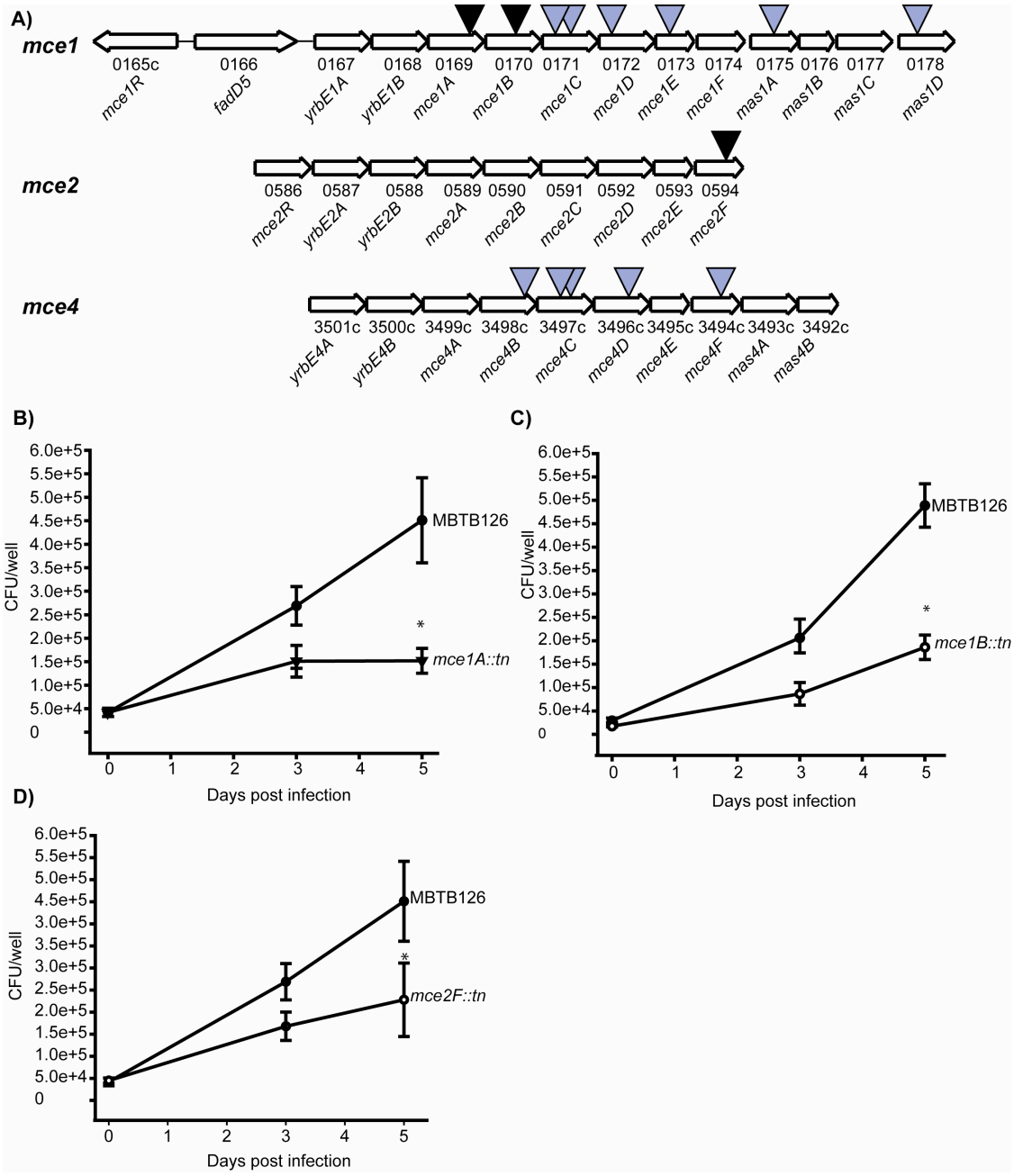

Figure 4.6. Some *M. tuberculosis mce::tn'blaTEM*-1 mutants have growth defects in macrophages, while others do not. **A)** Representation of the *mce1, mce2* and *mce4* operons in *M. tuberculosis* H37Rv. Black triangles indicate TnTEM insertions that resulted in macrophage growth defects. Grey triangles mark TnTEM insertions that did not lead to attenuation. **B-D)** *M. tuberculosis* MBTB126 or B) *mce1A::tn'blaTEM-1*, C) *mce1B tn'blaTEM-1*, or D) *mce2F::tn'blaTEM-1* were used to infect bone marrow derived macrophages. At 4 hours (Day 0), 3 and 5 days post infection, macrophages were lysed and plated to enumerate intracellular bacteria. Error bars indicate standard deviation of the mean. *, $p \leq 0.05$. Each graph shows representative data from 4 independent experiments.

are essentially competition assays. Subtle defects may be exaggerated during co-infections; phenotypic defects may not be evident when mutants are examined individually.Our mutant with a Tn'blaTEM-1 insertion in *mce2F* was also attenuated for intracellular growth (Fig. 4.6). To our knowledge there is no prior demonstration of a role in macrophage growth for the *mce2* operon. In fact, unlike *mce1, mce3,* and *mce4*, expression of the *mce2* operon has not previously been observed under any infection conditions tested (52). The fact that we recovered an active Tn'blaTEM-1 insertion in *mce2F* indicates that in our plate-grown *M. tuberculosis* this gene is expressed. Moreover, the intracellular growth phenotype we observed for the Tn'blaTEM-1 insertion mutant in *mce2F* not only indicates the importance of Mce2 to growth inmacrophages but (Fig. 4.6D) also indicates Mce2F expression by *M. tuberculosis* during intracellular growth.

Unlike our results of testing downstream mutations in *mce1*¸ in *mce2* the attenuated mutant we obtained has a transposon insertion in the last gene in the *mce* operon. Because the *mce2F* gene is the final gene in the *mce2* operon, it is unlikely that the phenotype we observed for this Tn'blaTEM-1 mutant is due to polar effects on downstream genes. Nonetheless, complementation experiments are still necessary to establish with certainty that the mutant phenotype of the *mce2F* transposon mutant is due to the insertion in *mce2F*. Such complementation experiments are currently underway. We also recovered several insertion mutants throughout the *mce4* operon (Fig. 4.6). However, none of these resulted in growth phenotypes in macrophages. The same was true in the TraSH in macrophages analysis of *mce4*, in that none were reproducibly underrepresented following macrophage infections with pools of mutants (77).

The role of the *mce* operons in *M. tuberculosis* remains unclear. As mentioned above, there might be a role in bacterial uptake by host cells. Mce1A facilitates uptake of latex beads by HeLa cells (6). Notably, Mce proteins Mce3A and Mce4A also promote uptake of coated beads or recombinant *E. coli* by HeLa cells (38, 86), but Mce2A does not. However, in respect to a potential role in cell entry it is noteworthy that in macrophage and in mouse studies with various *mce* mutants there is no suggestion Mce functions in initial infection or macrophage uptake (48, 90). We also did not see any differences between MBTB126 and the various *mce::tn'blaTEM-1* mutants when comparing macrophage uptake of each strain. However, it remains a possibility that redundancy between *mce* operons is masking phenotypes of individual mutants.
The presence of two putative ABC transporter permeases in each operon has led to a separate hypothesis that each *mce* operon encodes an ABC transporter apparatus (21, 48, 98). ABC transporters are protein complexes that function in the import or export of a wide-variety of molecules across membranes (12). ABC transporters sometimes rely on surface-localized substrate binding proteins to import substances. Interestingly, Mce proteins share homology with substrate binding domains of proteins involved in ABC transport (21, 53).

It is possible that the Mce proteins work together to transport lipids (21, 105). A study that described a screen for genes that interact with *mce* family members identified many genes with predicted functions in lipid transport or modification as potential interactors (48). Additionally, the *mce1* operon includes an ORF encoding FadD5, a protein with homology to fatty acid synthases. Further suggestion that Mce proteins may have a lipid related function comes from experiments performed with *Rhodococcus*, a

soil actinomycete related to *M. tuberculosis*. In a study of cholesterol catabolism in *Rhodoccocus* a region of the genome that shares homology with genes in *M. tuberculosis mce4* was identified (105). When the *Rhodococcus mce4* homologs are deleted, the mutant strain is no longer able to grow on cholesterol indicating a function in cholesterol transport or catabolism (68, 105). Thus, it is tempting to speculate that the role of at least *mce4* in *M. tuberculosis* may be in cholesterol uptake. Cholesterol is important for phagocytosis of mycobacteria and inhibition of phagosome maturation (31, 42, 73).

Rv0199

Considered to be a "core mycobacterial gene" (63), Rv0199 is found in all *Mycobacterium* genomes sequenced thus far, but virtually nowhere else. While there is very little homology to inform on its function, it is interesting to note that Rv0199 is 41% similar to the *mce*-associated protein Rv1972 encoded at the end of the *mce3* operon.

The protein encoded by Rv0199 is predicted to have a single TM domain proximal to the N-terminus of the protein (AA 42-64), indicating that the majority of the protein is located in the cell envelope. The Tn'blaTEM-1 insertion is in the middle of the ORF (Fig. 4.7A). Consistent with our demonstration that an in-frame insertion of Tn'blaTEM-1 in Rv0199 resulted in a macrophage growth defect (Fig. 4.7B), this ORF was predicted to play a role in virulence by two independent screens for mutants attenuated for growth in mice and macrophages. The first was the TraSH analysis of *M. tuberculosis* mutants in mice that showed a Rv0199 transposon mutant is underrepresented in output pools following infection (89). The second was screening of pooled *M. bovis* BCG transposon mutants for attenuated phenotypes in cultured

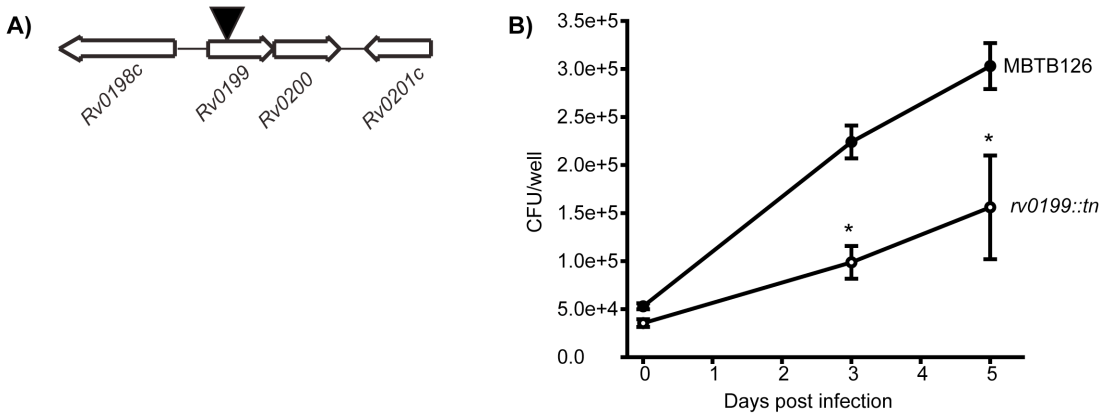

Figure 4.7. ***M. tuberculosis rv0199*::*tn'blaTEM-1*** **has a growth defect in macrophages A)** Organization of genes surrounding Tn'blaTEM-1 insertion (triangle) in Rv0199**. B)** Macrophages were infected with *rv0199::tn'blaTEM-1*, and intracellular growth of bacilli was monitored at the time points indicated as described for Fig. 4.6. Error bars indicated standard deviation of the mean. Graph shows representative data from 7 experiments. *, $p \leq 0.01$

macrophages. This screen identified transposon mutants in genes surrounding and including Rv0199 as underrepresented in mutant pools grown in macrophages over 72 hours (92). Because this region, spanning from Rv0199 to Rv0208c, codes for several predicted membrane proteins that are involved in pathogenesis, it was renamed the *v*irulence *a*ssociated *m*embrane *p*rotein region, or VAMP (92) for intracellular growth following macrophage infection with transposon mutant pools (77). In comparison to other studies, our study is the first to test a transposon mutation in *rv0199* on an individual basis. We can not rule out the possibility that the macrophage growth defect we observed for *rv0199::tn'blaTEM-1* is due to polar effects on downstream genes. The stop codon for Rv0199 overlaps the start codon of Rv0200, another ORF with no homology to any gene outside of mycobacteria. Rv0200 is also implicated as important for growth in macrophages in pooled infections (92) (Fig. 4.7A). The potential virulence roles for the genes surrounding *rv0199*, as well as the likelihood that Rv0199 is co-transcribed with its downstream neighbor, make it imperative that we perform complementation assays to determine if the lack of *rv0199* function is indeed the cause of the macrophage growth phenotype.

MmpS5

M. tuberculosis encodes several *m*ycobacterial *m*embrane *p*roteins, a protein family restricted to mycobacteria and a few closely related genera. Members of the Mmp family are divided into two groups – small, or MmpS, and large, or MmpL. MmpS5 (Rv0677c) is one of four MmpS proteins in the *M. tuberculosis* genome. Each MmpS protein is small, approximately 140 amino acids in length, and shares 59-70% similarity

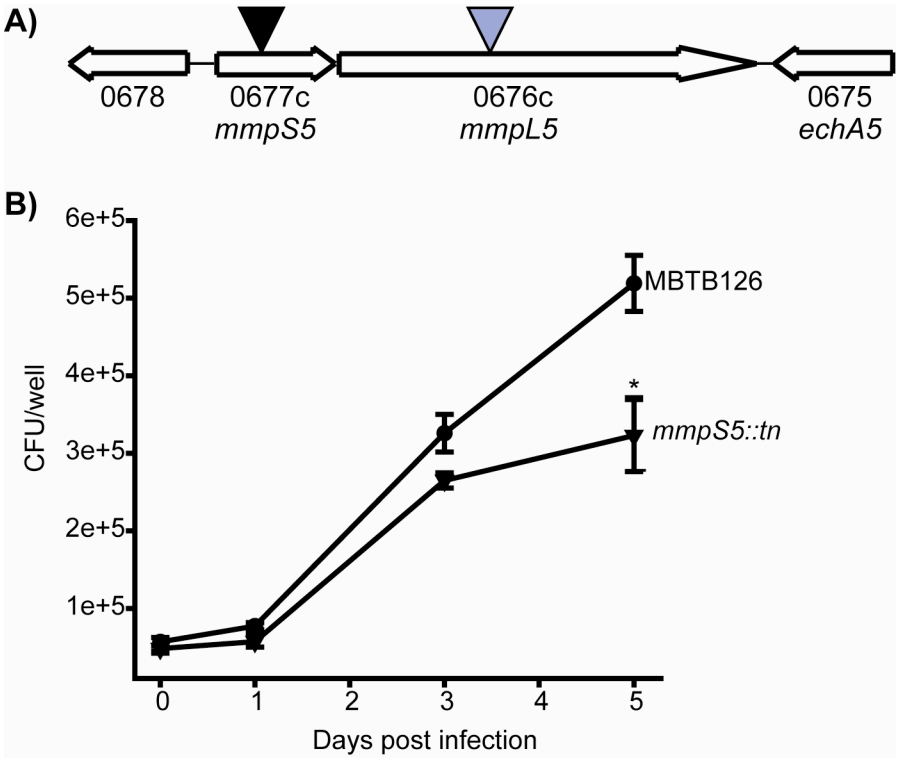

Figure 4.8. ***M. tuberculosis mmpS5::tn'blaTEM-1*** **has a growth defect in macrophages, but a** ***mmpL5 ::tn'blaTEM-1*** **insertion mutant does not. A)** Relative location of *mmpS5* in the *M. tuberculosis* H37Rv genome. The black triangle indicates a Tn'blaTEM-1 insertion that resulted in a macrophage growth defect, and the grey arrowhead indicates a Tn'blaTEM-1 insertion that did not lead to attenuation. **B)** Bone marrow derived macrophages were infected as described for Fig 4.6. Error bars represent standard deviation of the mean. *, p=0.01. The graph shows representative data from one of five independent experiments.

with the other MmpS family members. Repeated testing of our *mmpS5::Tn'blaTEM-1* mutant in murine bone marrow macrophages revealed an intracellular growth defect (Fig. 4.8). While its name implies that it is a membrane protein, our bioinformatic analysis revealed a cleavable signal sequence, as opposed to a TM domain at the beginning of MmpS5. The probability that MmpS5 has a signal peptide is 0.921, well above the threshold value of 0.5 for the prediction algorithm (9). Further, MmpS5 was identified as a protein secreted into the culture filtrate of *M. tuberculosis* (60). Thus, MmpS5 may be mislabeled as a membrane protein, but not necessarily misannotated as an MmpS protein due to its strong homology with the three other MmpS family members in *M. tuberculosis*.

No direct investigation of function or role in virulence has been conducted for any of the MmpS proteins. However, in all cases the *mmpS* genes are placed directly upstream of genes that code for MmpL proteins. Therefore, it is likely that each *mmpS-L* pair is co-transcribed, although co-transcription has not been experimentally shown. It is hypothesized that the functions of MmpS and MmpL proteins are linked. Beyond screens for phenotypes in pools of transposon mutants (55, 77, 89), there have been further direct studies of MmpL proteins, but not of MmpS proteins, in pathogenesis (34).

MmpL proteins are predicted to be part of the Resistance, Nodulation and Division (RND) family of molecules. MmpL proteins typically have 12 TM domains with two large domains predicted to reside on the extracytoplasmic face of the bacterial membrane. This predicted toplogy is supported by the sites of insertion we identified for 'BlaTEM-1 and TnPhoA reporters in MmpL4 (16), Fig. 4.4). In other bacteria, some RND family members function as efflux pumps (74, 104). The best-characterized efflux

pumps of the RND family are in Gram-negative bacteria, and usually rely on cooperative function of an associated outer membrane canal protein (OMP) and a membrane fusion protein (MFP). One example of such a complex is the AcrA-AcrB-TolC efflux system in *E. coli*. AcrB is the RND family member and an integral membrane protein. AcrA acts as the MFP, and is found in the periplasm. The OMP is TolC, which is an outer membrane protein. The coordinated action of the RND pump with the MFP and OMP proteins allows the pump to expel compounds directly to the outside of the bacterial cell, bypassing the periplasmic space (37, 74). It has been suggested that members of the MmpL family share similarity with the AcrB transporter protein (33, 34). In *E. coli*, AcrA (the MFP) stimulates the activity of the AcrB pump (3, 70). One possibility for a function of the MmpS proteins is a role in activation of MmpL pumps. However, at this time such a function is pure speculation as there is little sequence homology between MmpS proteins and the activating proteins of Gram-negative pumps.

There are 12 MmpL proteins in *M. tuberculosis* H37Rv. Insertional inactivation of 11 of these showed that 4 are required for pathogenesis in mice, while MmpL3 is predicted to be essential for in vitro growth, as attempts to inactivate it have been unsuccessful (34, 88). Some *mmpL* genes, as is the case for *mmpL7* and *mmpL8*, are surrounded by lipid synthesis and export genes, so it was predicted that these MmpL proteins are involved in transport of lipids beyond the cytoplasm. This turns out to be true. MmpL7 is required for export of phthiocercol dimycocersosate (PDIM) and MmpL8 is required for export of the sulfolipid SL-1 (19, 27, 28, 35). Unfortunately, these studies do not shed light on the function of MmpS proteins in *M. tuberculosis*. Neither MmpL7 nor MmpL8 are associated with an MmpS protein counterpart. If the

functions of MmpS and MmpL are indeed coupled for transport or efflux, the four MmpS proteins may work with more than one of the 12 different MmpL transporters in *M. tuberculosis*. We recovered Tn'blaTEM-1 insertions in *mmpL1*, *mmpL4, mmpS2, mmpS5, mmpL5 and mmpL9* (Supplemental Table 4.1, Fig. 4.4), but, aside from *mmpS5::tn'blaTEM-1*, none of these led to a significant growth defect in macrophages.

A recent study suggests a possible efflux function for MmpS5. While testing the potential use of azoles in treatment of mycobacteria, it was shown that *mmpS5* and *mmpL5* may code for a complex that contributes to azole resistance. Strains spontaneously resistant to azole drug over-express *mmpS5* and *mmpL5* (67). It is not yet tested whether *M. tuberculosis* strains that lack MmpS5 and/or MmpL5 are more sensitive to killing by azoles.

Until we perform complementation, it remains possible that the insertion in *mmpS5* is having polar effects on the expression of *mmpL5*, as the coding sequences of the two genes overlap. However, a transposon inserted in the *mmpL5* gene that truncated its final 508 amino acids did not result in an intracellular growth phenotype for this mutant, which argues against this possibility (Fig. 8A). It is still necessary to perform complementation experiments for *mmpS5::tn'blaTEM-1* and these experiments are in progress.

CtaC

Like other bacteria, mycobacteria produce energy by establishing a proton gradient across the cytoplasmic membrane. This gradient is generated as electrons flow down the electron transport chain (ETC). Electrons released from an initial donor pass

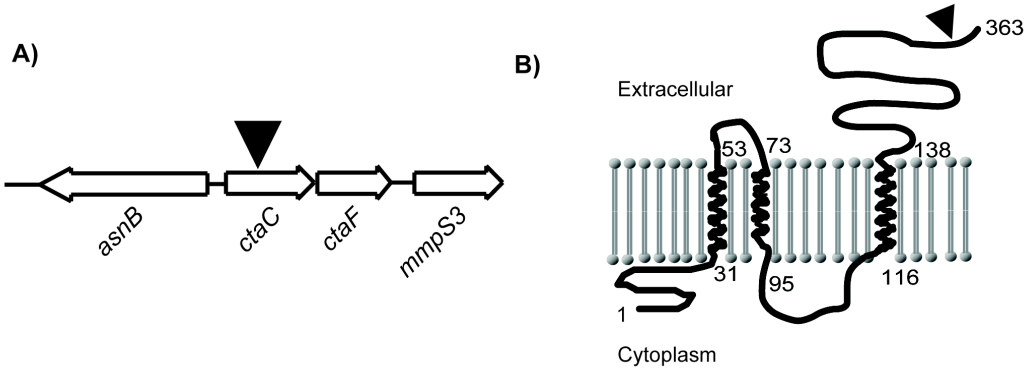

Figure 4.9. Location of *ctaC::tn'blaTEM-1* insertion, and predicted topology of CtaC. A) The relative location of *ctaC* in the *M. tuberculosis* genome is shown. The black triangle indicates the site of Tn'blaTEM-1 insertion. **B)** Membrane topology of CtaC as predicted by the TMHMM algorithm (91). The site of 'BlaTEM-1 fusion is indicated with a triangle.

through a series of alternating oxidized and reduced flavoproteins, quinones and cytochromes. This process terminates, in the case of aerobic respiration, in the reduction of O_2 to H_2O. During this process, protons are transported to the extracytoplasmic side of the membrane, establishing an energy gradient that can be harnessed for the generation of ATP. In silico analyses suggest that mycobacteria contain multiple terminal respiratory oxidases, one of which is the four-subunit cytochrome c oxidase complex responsible for catalyzing the reduction of oxygen when O_2 is sufficiently available (50, 66, 78).

The *ctaC* gene encodes the membrane bound subunit II of the cytochrome c oxidase. It has three predicted TM domains (91), one of which resides in a possible cleaved export signal sequence in the N-terminus (9). CtaC protein is found in culture filtrate, despite the strongly predicted TM domains, indicating that some protein is shed into the medium during in vitro growth (60). We can combine the information provided by TM and signal sequence prediction algorithms with the sites of our active Tn'blaTEM-1 insertions to predict the protein topology of CtaC. We predict that the N-terminus is on the cytoplasmic side of the membrane followed by three TM domains (Fig. 4.9B). The cytochrome oxidase domain is predicted to span from AA 116 to AA 322. Therefore the active domain of CtaC is beyond the extracyotoplasmic face of the membrane, and its orientation is similar to that predicted for *Bacillus subtilis* CtaC (11).

CtaC is predicted to be essential in *M. tuberculosis* H37Rv and an attempt to delete the gene is *M. tuberculosis* is reported to have been unsuccessful (64, 88). It is likely that we were able to obtain our Tn'blaTEM-1 insertion in *ctaC* because the majority of the functional protein remains unaltered. Our insertion is in the extreme C-

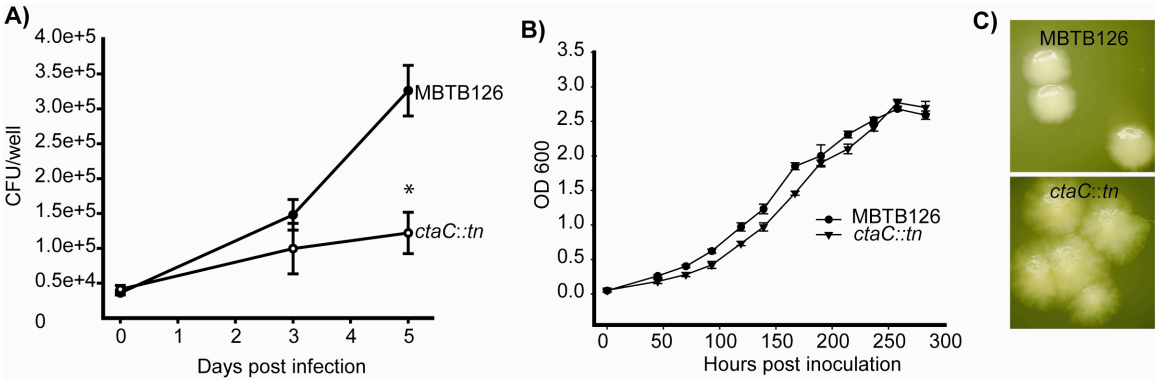

Figure 4.10. ***M. tuberculosis ctaC::tn'blaTEM-1*** **has a growth defect in macrophages but not in broth culture, and has a colony phenotype on agar plates**. **A)** Macrophage infection done as described for Fig. 6. Error bars indicate standard deviation of the mean. *, p = 0.02. Graph shows representative data from three independent experiments. **B)** *M. tuberculosis* MBTB126 or *ctaC::tn'blaTEM-1* were inoculated to OD 0.1 in 7H9 supplemented with ADS, glycerol, tween and hygromycin. Growth of triplicate cultures of each strain was monitored by optical density at OD600 for the time period indicated. Graph shows representative data from five independent experiments. **C)** *M. tuberculosis* MBTB126 or *ctaC::tn'blaTEM-1* were grown to mid-exponential phase in 7H9 and spread on 7H10 agar plates. Colonies were examined after an incubation period of 25 days at 37º.

terminus of the protein and would truncate the final 13 amino acids. This insertion mutant had a reproducible growth defect in macrophages in our experiments (Fig 4.10A).

Unlike the attempts made in *M. tuberculosis*, a *ctaC* mutation in the non-pathogenic *M. smegmatis* is viable, although the mutant has a growth defect in 7H9 medium (64). To ensure that the macrophage growth defect we observed was not simply due to a lack of overall fitness, we compared the growth rate of the *ctaC::tnTEM-1* mutant strain to MBTB126 and saw no difference (Fig. 4.10B). A colony phenotype was also observed for the *ctaC::tn'blaTEM-1* mutation when grown on 7H10 agar supplemented with tween. While the MBTB126 CFU were compact and slightly rough around the edges, the *ctaC::tn'blaTEM-1* mutant CFU were very rough, spread out and varied in size (Fig. 4.10C).

As mentioned above, we observed that our *ctaC::tn'blaTEM-1* mutant had an intracellular growth defect when compared to MBTB126. One explanation for this phenotype is that the decreased intracellular growth of the *ctaC::tn'blaTEM-1* mutant results from compromised CtaC activity and reduced establishment of a membrane potential for the generation of ATP. Another possibility is that the intracellular growth defect of *ctaC::tn'blaTEM-1* may be caused by polar effects of this mutation, and not the insertion in *ctaC* itself (Fig. 4.9A). The *ctaF* gene is directly downstream of *ctaC*. CtaF shares high homology with cytochrome c subunit proteins in other mycobacteria and *Rhodococcus*, and is highly likely to be co-transcribed with *ctaC* because its translational start codon is only 7 nucleotides from the translational stop of *ctaC* (1, 4, 26). It is possible that CtaF works in concert with CtaC and contributes to the structure of the multi-subunit cytochrome c oxidase in mycobacteria, or it may function in another

capacity. Complementation experiments that restore CtaC expression to the *ctaC::tn'blaTEM-1* mutant are required to make any meaningful conclusions about the role of CtaC in *M. tuberculosis* intracellular growth.

PpsB, DrrC and LppX

Pathogenic mycobacteria have unusually thick cell envelopes composed of unique lipids (51). One of these unique outermost surface lipids is phthiocerol dimycocerosate (PDIM) (7, 71). In numerous studies, *M. tuberculosis* mutants that do not make or do not properly localize PDIM to the outer surface of the bacterium are attenuated for virulence. The majority of the genes responsible for synthesizing, modifying and exporting PDIM in *M. tuberculosis* are found in neighboring clusters in the genome. PpsA-E are encoded within a single poly-cistronic message that starts with *fadD26* (Rv2930), continues with the *drrABC* locus and ends with *papA5* (18, 26). The next gene in the PDIM genomic region codes for mycocerosic acid synthase (*mas*), which is followed by the co-transcribed *fadD28* and *mmpL7* genes (18). The final gene in the PDIM cluster encodes the lipoprotein LppX (Fig. 4.11A).

We identified three carbenicillin resistant clones with Tn'blaTEM-1 insertions in genes known to be required for PDIM synthesis or localization: *ppsB, drrC and lppX*. (18, 19, 28) (Fig. 4.11A). Both the *ppsB* and *lppX* Tn'blaTEM-1 insertion mutants showed intracellular defects when compared to MBTB126 (Fig. 4.11). To our knowledge, this is the first direct testing of *ppsB* or *lppX* mutants for intracellular growth in macrophages. Further, TraSH did not predict them to be important for growth in macrophages (77). A published insertion mutant in the third gene that codes for the 6-

transmembrane domain protein DrrC, was previously tested and reported to be defective for growth in macrophages (18). However, our *drrC::tn'blaTEM-1* mutant, with an insertion truncating the protein at amino acid 57 (Supplemental Table 4.1) showed no growth defect in macrophages in 4 independent experiments (Fig. 4.11D). Differences in *M. tuberculosis* strain backgrounds or macrophage assays may have led to the discrepancy between our observed data and published results.

We have clues to the function of most of the gene products encoded in the PDIM region (44). Each Pps protein is a type I modular polyketide synthases responsible for turning C_{18+} fatty acids generated by fatty acid synthase I into phthiocerol, the precursor to the *M. tuberculosis* PDIM. PpsA-E take the fatty acid initially activated and supplied by FadD26 (102, 103) and elongate it by adding saturated alkyl groups in a stepwise manner. PpsE makes the final modification, which serves to complete the synthesis of phthiocerol (18, 103). In the meantime, Mas generates mycocerosic acids from modified C_{12}-C_{18} fatty acids (7, 103). These mycocerosic acids become linked to the Pps-modified acyl chains of phthiocerol, resulting in phthiocerol dimycocersosate (PDIM),and PapA5 is an acyltransferase that is responsible for this linkage (71, 103).

Once the complex lipids are constructed, DrrC, MmpL7 and LppX are required for proper transport of PDIM across the cytoplasmic membrane to its final destination at the bacterial surface. DrrABC are predicted to form an ATP-binding cassette (ABC) type of transporter (30), where DrrA is thought to be the ATP binding component and DrrB or DrrC is the likely integral membrane protein that transports substrates beyond the cytoplasmic membrane. A DrrC insertion mutant is able to make PDIM, but unable to

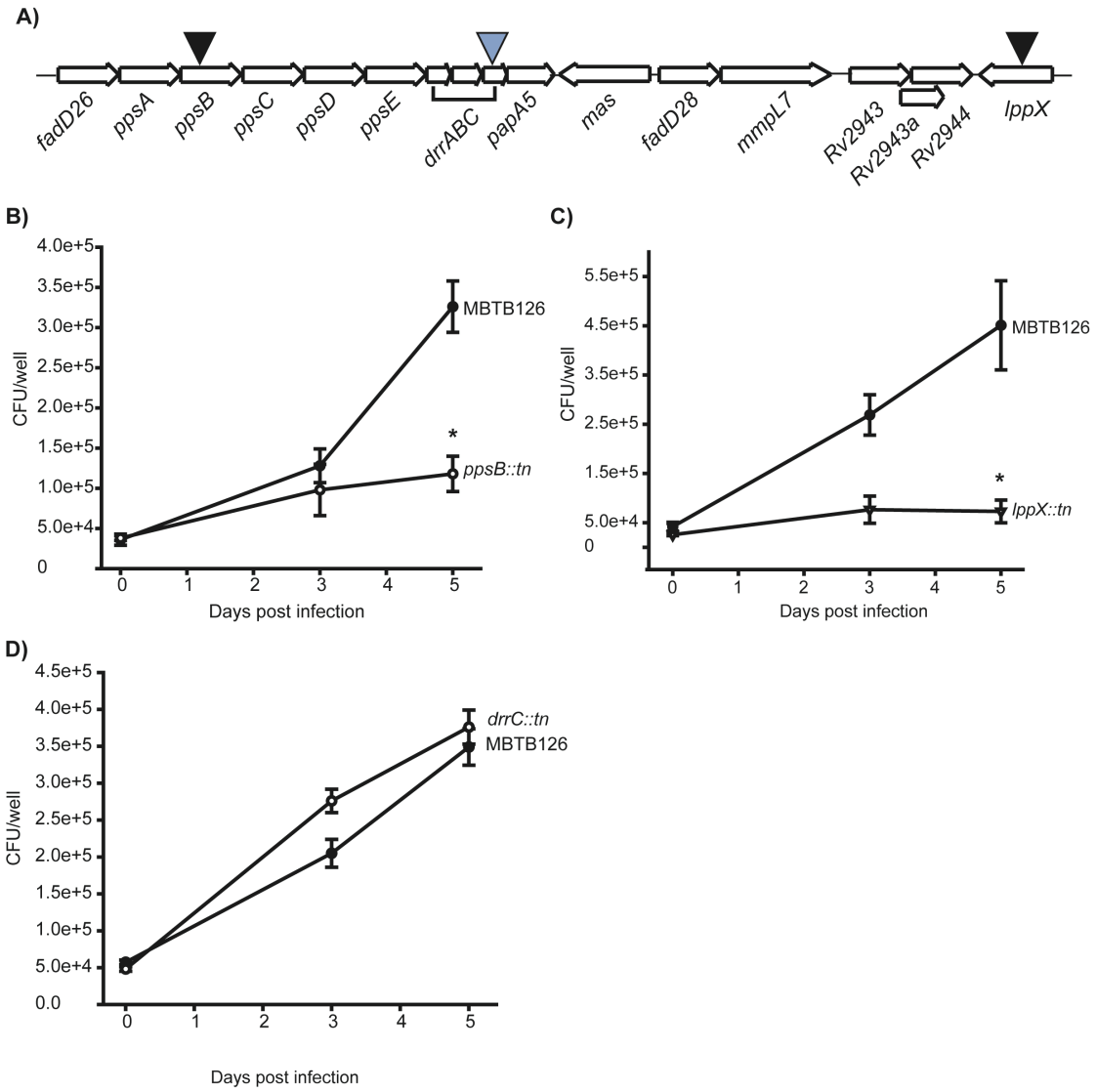

Figure 4.11. ***M. tuberculosis ppsB::tn'blaTEM-1*** **and** ***lppX::tn'blaTEM-1*** **have intracellular growth defects, but the** ***drrC::tn'blaTEM-1*** **mutant does not. A)** Representation of insertions in the PDIM production and transport locus of *M. tuberculosis*. Black arrowheads represent Tn'blaTEM-1 insertions that resulted in attenuation in macrophages. The gray arrowhead indicates a Tn'blaTEM-1 insertion that did not lead to attenuation. Rv2943-2944 encode for potential transposase genes from an insertion element. Rv2943a possibly contains a frame-shift mutation. **B-D**) Murine bone marrow-derived macrophages infected with *M. tuberculosis* MBTB126 or **B)** *ppsB:tn'blaTEM-1*, **C)** *lppX::tn'blaTEM-1*, or **D)** *drrC::tn'blaTEM-1*. Infections were done as described in Fig. 4.6. Error bars indicate standard deviation of the mean. *, p <0.01. Each graph shows representative data from one of 3 or 4 independent replicates.

properly localize it to the cell wall (18). The polytopic MmpL7 membrane protein is also required for export of PDIM to the cell wall. An insertion mutant in *mmpL7* makes PDIM but does not localize it to the cell surface (18, 28). Finally, the lipoprotein LppX is responsible for the localization of PDIM to the outermost bacterial surface (57, 60, 93). When the PDIM molecules of *lppX* mutants were examined, they are unaltered when compared to PDIM extracts obtained from WT *M. tuberculosis*. However, it was discovered that, unlike WT bacteria, the *lppX* insertion mutant does not shed PDIM molecules into the culture medium (93). The LppX crystral structure revealed that this mycobacterial protein resembles *E. coli* proteins LolA and LolB, both of which are required for localization of lipoproteins at the outer membrane (93, 95, 96). LolA is active in the periplasm of *E. coli*, acting as a chaperone to transfer lipoproteins to the outer membrane lipoprotein localization factor LolB (101). The predicted structures of LppX, LolA and LolB share a hydrophobic cavity where lipids are predicted to bind. C_{16}-C_{18} fatty acids do bind to purified LppX, and virtual modeling demonstrates that the hydrophobic cavity is the appropriate size to accommodate a PDIM molecule (93).

The active Tn'blaTEM-1 insertion we identified in *ppsB* serves to demonstrate that PpsB is an exported protein. To our knowledge, this is the first evidence of PpsB being an exported protein in *M. tuberculosis*. TM domain prediction servers differ on the existence and number (0-6) of TM domains PpsB contains (46, 91). In a large-scale proteomics study, PpsE is reported in the membrane fraction of *M. tuberculosis* (108). As for PpsB, our bioinformatic analysis of the PpsE amino acid sequence predicted 0-5 transmembrane segments (46, 91). The *ppsB::tn'blaTEM-1* mutant carries an insertion following AA 613. This site of insertion is at the beginning of the predicted acyl

transferase domain of the protein that is likely responsible for covalently acylating fatty acids extended initially by PpsA (103). The ß-lactam resistance resulting from our Tn'blaTEM-1 insertion in *ppsB* indicates that at least a small relative portion of the predicted acyl transferase domain faces the extracytoplasmic face of the bacterial membrane.

While we are the first to report that *lppX* and *ppsB* transposon mutants exhibited intracellular growth defects, these results were not surprising. It is well established that mutants with defects in PDIM synthesis or transport are attenuated for virulence. Examination of other data surrounding the host reaction to infection with PDIM-less mutants gives clues as to why these mutants are defective for intracellular growth. Insertions in *fadD26*, *mmpL7*, or *drrC* all result in attenuation when compared to the intracellular growth of WT bacteria (19, 81). *M. tuberculosis* mutants that lack PDIM are more sensitive than WT to RNI and SDS in vitro, suggesting that the PDIM outer layer may act as a permeability barrier for the bacterium (18). Another aspect of PDIM's role in pathogenesis is its interaction with host innate immune response elements. The PDIM-less mutants induce more production of the pro-inflammatory cytokines TNFα and IL-6 than WT *M. tuberculosis* in infected macrophages. This was true for both the *fadD26* and *mmpL7* insertion mutants, indicating that having PDIM and having it properly localized are equally important for dampening a pro-inflammatory response (81). It is possible that the intracellular and in vivo growth defects of PDIM mutants stem from the inability to dampen the immune response and adequately resist RNI or other stresses produced by macrophages.

Whole animal aerosol infections with Tn'blaTEM-1 mutant strains

Mice are often used as a model system for tuberculosis disease. Following inhalation of 100-200 *M. tuberculosis* bacilli into the lungs, one of the first events observed is bacterial replication in lung alveolar macrophages. Rapid growth and dissemination of the bacteria occurs 14 - 21 days following infection, at which point, the adaptive immune system curtails further growth. Lymphocytes that specifically recognize the bacteria sequester intracellular *M. tuberculosis* in granuloma-like lesions, and bacterial populations persist at a steady state. Mouse strains that are relatively resistant to infection, including the C57BL/6 mice used in our studies, typically succumb to disease between 150 and 200 days post infection (15, 43).

M. tuberculosis must be taken up by and replicate in lung macrophages to establish infection following aerosol exposure. For this reason, we hypothesized that our Tn'blaTEM-1 mutants with macrophage growth defects would also be defective in a whole-animal model of tuberculosis infection. For two of our eight mutants, the process of testing them in a mouse model of tuberculosis infection has begun but is not yet complete. The mutants chosen for the initial evaluation have transposon insertions in *mmps5* or *ppsB*. These mutants were chosen, in part, because there is no prior direct testing in mice of *M. tuberculosis* mutants with mutations in these genes. Each of the strains was delivered by aerosol to C57BL6 mice. One day post-aerosol infection, four mice from each infected group were euthanized. The lungs of the sacrificed animals were homogenized and plated to recover viable CFU to measure the initial dose of bacteria received. At later time points, the bacterial burden in lungs, liver, and spleen was determined in groups of animals. This allowed growth and dissemination of *M.*

tuberculosis to be followed throughout the course of each infection. Long-term survival in sets of animals was also followed. In these animal experiments, mice were infected in parallel with MBTB126, which was used as the parental control strain, for comparison.

MmpS5

Mice infected in parallel with the *mmpS5::Tn'blaTEM-1* mutant or MBTB126 were evaluated for bacterial burden in organs over time and for length of long-term survival. No differences were observed in the growth or persistence of the *mmps5* mutant in lungs at multiple time points post-infection (Fig. 4.12A). There was also no difference observed in the ability of the *mmps5* mutant to disseminate to liver and spleen or to grow in these organs over time (data not shown). Furthermore, there was no difference in the length of survival of animals infected with the *mmpS5::Tn'blaTEM-1* mutant versus those infected with MBTB126 (Fig. 4.12B).

While it is important to repeat this experiment, initial results suggest that *mmpS5* has a reproducible growth defect in primary murine macrophages but does not have a growth defect in the organs of C57Bl/6 mice following a low-dose aerosol infection. An *mmps5* mutant has not previously been directly tested in macrophages or mice, but has been assayed in the pool of transposon mutants evaluated in the TRaSH studies, which allows prediction of the possible contribution the *mmpS5* gene plays in mice and macrophages (77, 89). While it seems odd that *mmpS5::tn'blaTEM-1* had an intracellular growth defect but no defect in the animal model of tuberculosis, our results are consistent with TraSH. The TraSH analysis of transposon mutants in pooled infections indicates that MmpS5 is important for growth in macrophages, but not in mice (77, 89).

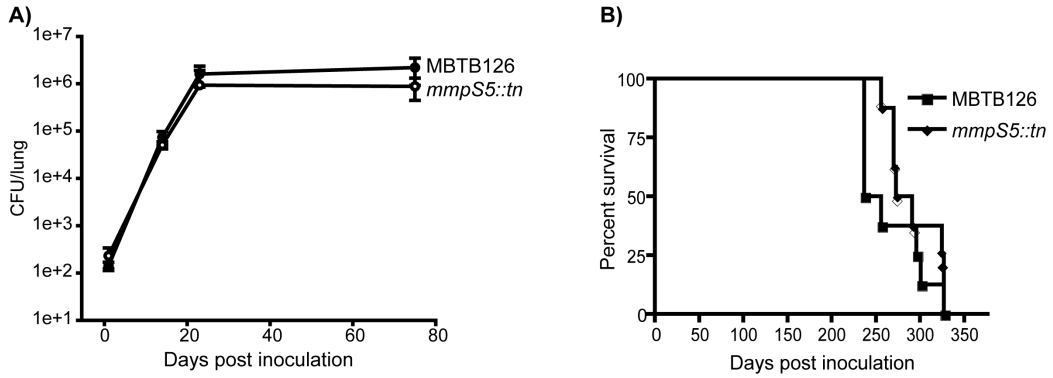

Figure 4.12. The *mmpS5::tn'blaTEM-1* mutant was not attenuated in mouse lungs, and did not affect survival of infected animals. C57Bl/6 mice were infected with the indicated *M. tuberculosis* strains by aerosol delivery. **A**) At 1, 14, 25 and 79 days post-infection, lungs of 4 infected animals per strain were homogenized, diluted and plated for resident CFU. Error bars indicate standard deviation. **B**) Eight infected mice per strain were monitored for survival over the period of days indicated.

PpsB

PpsB is one of the polyketide synthases responsible for extending lipid molecules on their way to becoming PDIM. While other mutations in the PDIM locus have been tested for their importance in whole animal models of tuberculosis (19, 28, 81), a *ppsB* mutant has not previously been tested. Our *Tn'blaTEM-1* insertion presumably disables the majority of PpsB (Fig 4.11A). We inoculated mice by aerosol with low doses of either the *ppsB::tn'blaTEM- 1* mutant or MBTB126, and compared the course of disease in each group of mice as described above (Fig. 4.13). We found that organ burdens of the mutant strain were one log lower over a 79-day long infection period than that of the parent strain of *M. tuberculosis.* Further, the mutant strain had approximately 1 log fewer bacteria disseminating from the lung to the liver and spleen than the parent strain (Fig. 4.13).

The importance of PDIM to pathogenesis was first shown in signature tagged mutagenesis experiments, where insertions in genes required for PDIM synthesis, modification or export result in mutants that are attenuated for growth in vivo (19, 28, 81). Insertion mutations in the *fadD26* promoter (which would presumably disrupt transcription of all downstream *pps* genes and therefore does not make PDIM) result in reduced growth in mouse lungs over 3 weeks (18, 19, 28). This *fadD26* mutant also has a defect in persisting in the lung (81). In our hands, the *ppsB* mutant strain did not have a similar persistence defect, as the bacterial populations in the lung remained at

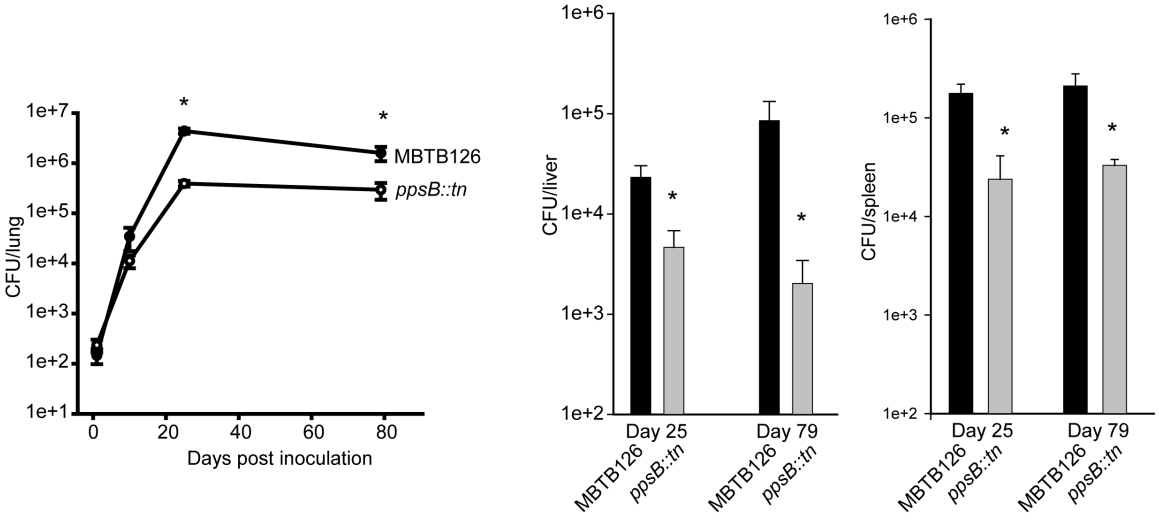

Figure 4.13. The *M. tuberculosis ppsB::tn'blaTEM-1* mutant was attenuated for growth in mouse lung, liver and spleen following infection by aerosol. C57Bl/6 mice were infected with approximately 10^2 CFU of *M. tuberculosis* MBTB126 or the *ppsB::tn'blaTEM-1* mutant. At 1 and 14 days post inoculation, lungs were removed from 4 infected animals per strain per time point. At 25 and 79 days, lungs, livers and spleens were removed. At all time points, dissected organs were homogenized, diluted and plated for CFU. Each time point represents data from 4 animals per strain. Error bars depict standard deviation. *, $p \leq 0.03$.

approximately 1 log lower CFU than WT. However, a persistence defect may have been observed if the organ burdens were determined at a later time point. Based on previous work of others, reviewed above, we suspect our *ppsB* transposon mutant is not synthesizing PDIM, and the lack of PDIM is the most likely cause of the defective intracellular and in vivo growth phenotypes we observed.

CONCLUSION

In this work, we constructed and utilized a new reporter transposon for large-scale identification of exported proteins in *M. tuberculosis*. By combining the β-lactamase reporter of protein export with transposon mutagenesis of *M. tuberculosis*, we were able to establish the subcellular localization of 118 unique *M. tuberculosis* proteins and simultaneously collect a library of mapped mutants with insertions in the genes encoding these exported proteins for subsequent analyses.

An important feature of our transposon-based reporter strategy is that it operates as a selection, not a screen. Transposon insertions that result in exported 'BlaTEM-1 fusions will grow in the presence of β-lactam antibiotics. Screening for an exported fusion phenotype is commonly required with the use of other reporters of protein export (16, 36, 58, 99) and is a much more labor-intensive approach. To screen over 80,000 Tn'PhoA insertions in the *M. tuberculosis* genome would have required at least 400 plates, each with approximately 200 CFU for an average of <1 CFU per plate expressing a construct capable of exporting 'PhoA. Further, rigorously screening over 80,000 transposon mutants individually for intracellular growth defects would have been impossible. However, a positive selection that allowed us to collect a library of mutants lacking exported proteins made screening for phenotypes on an individual scale feasible.

Interestingly, all false positive carbenicillin resistant clones were obtained while troubleshooting our selection strategy. Once our selection concentration of carbenicillin was optimized at 20 μg/ml, 100% of the carbenicillin resistant clones obtained had '*blaTEM-1* in frame with a gene predicted to encode an exported protein. There was a potential to identify mutants that became spontaneously resistant to β-lactams due to Tn'blaTEM-1 insertion in genes that code for cell wall synthesis or maintenance components, as formation of the bacterial cell wall is the target of β-lactam attack (109). However, we did not identify any transposon mutants that fell into this category.

Our β-lactamase fusions provided us with a powerful method to gain experimental evidence of the exported nature of individual proteins. The exported fusions also helped to define the extracellular domains of integral membrane proteins. We obtained multiple insertions in some proteins with more than one TM domain. By combining the information from algorithms that predict TM domains with the site of Tn'blaTEM-1 insertion, we could predict the location of extracellular domains for integral membrane proteins (Fig. 4.4). The mapping of extracytosolic domains will be particularly useful in studying the many membrane proteins we identified that have no known function.

Our first objective was to comprehensively identify the exported proteins of *M. tuberculosis.* However, as with every strategy that uses whole genome approaches, we were not able to collect insertion mutants in every protein exported by the bacterium. Using a calculation based on a Poisson distribution (45), we calculate that our screen is not yet saturated, and one in every five carbenicillin resistant Tn'blaTEM-1 insertions identified in future rounds of selection will be in an exported ORF we have not yet

identified. Also, because transposon insertion is a pre-requisite for our identification strategy, we could not select for exported proteins that are essential for growth – the genes that encode this class are predicted to make up approximately 20% of the genome (26, 88). We also could only select for exported fusions that were expressed and exported under the experimental conditions used – growth on 7H10 agar with 20 μg per ml of carbenicillin at 37 ºC for 3-4 weeks. Additionally, we hypothesize that there is a suite of proteins exclusively expressed or exported when *M. tuberculosis* resides in the host. Such proteins would be missed in our plate-based selections. We are currently developing strategies to identify this subset of exported proteins (See Chapter 3).

There are also likely to be several genes encoding exported proteins that, even if receiving an in-frame insertion, cannot code for active fusions because their products are not exported if fused to 'BlaTEM-1. The attachment of the β-lactamase moiety might block interactions with export machinery, or create a fusion that is too large to be exported. It is also possible that 'BlaTEM-1 fusions to specialized export substrates are indeed exported, but not functional because they are not properly folded. We had hypothesized that proteins exported by specialized transport pathways in *M. tuberculosis* could be discovered by our Tn'blaTEM-1 reporter system. This β-lactamase reporter has been used in the past to report on the location of proteins exported by the Type III (23, 62) and Type IV (32) specialized pathways in other pathogens. Type III and IV secreted substrates lack classical signal sequences. However, every exported 'BlaTEM-1 fusion we identified was to a protein with a predicted Sec or Tat signal sequence or TM domain (Table 4.4 and Supplemental Table 4.1). *M. tuberculosis* encodes at least two specialized export systems: SecA2 and ESX-1. These export pathways are required for *M.*

tuberculosis pathogenesis, and both export a small subset of proteins that lack N-terminal signal sequences (2, 79). To test whether the specialized systems of mycobacteria were capable of exporting substrates fused to 'BlaTEM-1, we directly fused the ESX-1 substrate ESAT-6 and the SecA2 substrate SodA to our β-lactamase reporter in different conformations. None of the resulting fusions conferred resistance to β-lactams on plates when expressed in *ΔblaC M. tuberculosis*, despite being expressed in the bacilli as shown by immunoblot (data not shown). Thus, we believe there are *M. tuberculosis* unconventional exported proteins that we will miss with our reporter transposon.

Notably, when we found multiple mutants with Tn'blaTEM-1 inserted in the same gene, there was a strong likelihood of seeing the insertion at the identical position. This was true in 27 out of 37 occurrences. For example, we identified six identical insertions in six independent phage infections in the ORF encoding Rv3802c. In fact, the majority of identical insertions were recovered from separate phage infections, which indicates they are not siblings but arose independently. The only known requirement for *Himar-1* insertion is a TA dinucleotide recognition sequence. We hypothesized that a potential reason for the recovery of repeat identical insertions was that not all TA sites in an ORF will allow 'BlaTEM-1 to be expressed in frame with the fused exported protein. However, when we made in silico insertions in every TA site in *rv3802c*, approximately half would have led to in frame expression of 'BlaTEM-1. We did obtain active 'BlaTEM-1 fusions along the length of proteins, thus there is no obvious problem with fusing mature domains of exported proteins to this reporter. However, a possible explanation for the identical repeat insertions is that the fused β-lactamase may, in some cases, inhibit proper folding of the fused protein which may block export. Conversely,

the fused domain may inhibit folding of ‘BlaTEM-1, causing the enzyme to be inactive. Therefore, for some exported proteins, only certain fusions may allow detectable export of ‘BlaTEM-1 following transposition. Inclusion of a flexible amino acid linker in front of and in frame with the reporter is a strategy we may use in designing similar reporter systems in the future. A linker may allow ‘BlaTEM-1 and its fused protein to fold independently of each other, which could lead to the identification of more exported protein fusions

We screened our library of transposon mutants in exported proteins for the ability to grow in macrophages. This identified eight transposon mutants with growth defects in macrophages over five days. For four of these mutants, CtaC, Mce2F, PpsB and LppX, it was shown for the first time that they were important for intracellular growth. Of the eight attenuated mutants, only PpsB, LppX and CtaC have a predicted function. Future experiments will focus on complementing the mutant phenotypes and determining the molecular roles these exported proteins have in virulence.

We hypothesized that transposon mutants with macrophage growth defects would have growth defects in the early phase of mouse lung infection, when macrophage replication is most evident (43). Of the two mutants that we tested in mice, the *ppsB* mutant proved to be defective for growth in both models while the *mmpS5* mutant did not. The different results between the two mutants in animals was surprising because the *ppsB* mutant was not more attenuated in macrophages than the *mmpS5* mutant; in fact, no Tn’blaTEM-1 mutant was statistically more attenuated for intracellular growth than any other (Fig. 4.5). We conclude that a growth phenotype in cultured macrophages (a cell-autonomous system) is not necessarily an indication that a particular mutant will be

unable to replicate in the more complex environment of the mouse lung. In reviewing the TraSH analysis data performed on *M. tuberculosis* mutants in macrophages and mice, there is also a subset of genes predicted as being required for growth in cultured macrophages that are not similarly predicted as being required in mice (77, 89). The *mmpS5* mutant is an example of one gene in this group and we obtained the same results in our testing of the *mmpS5::tn'blaTEM-1* mutant in macrophages and mice (Figs. 4.8B and 4.12A). It was proposed that this difference seen in the TraSH analysis may be due to the fact that TraSH mutants attenuated in mice were evaluated in spleen, not in lung, where macrophage replication is occurring initial exposure via the inhalation route (77). However, we evaluated our *mmpS5* mutant in macrophages and in lung, and saw no attenuation in vivo. There are also genes predicted by the TraSH studies as being defective in mice but not in macrophages – insertions in the *mce4* operon fall into this category. These mutants may represent ORFs that are not required in the early stages of infection, but are necessary for dissemination and persistence of the bacteria within the host. Our results led us to conclude that we cannot assume that attenuation in cultured cells will translate into attenuation in vivo, and vice versa. We hypothesize that the attenuation of particular mutants in macrophages may be masked by complex host-mycobacterial interactions in vivo. Therefore, we can use cell assays to identify macrophage responses that are specific to individual *M. tuberculosis* mutants. The characterization of phenotypes witnessed in small-scale macrophage culture will help our understanding of *M. tuberculosis* pathogenesis.

Our screen for Tn'blaTEM-1 mutants with intracellular phenotypes complements the TraSH saturating mutagenesis infection assays performed in macrophages (77). For

those mutants identified in both assays, our individual infections provide more detailed information about mutant phenotypes. However, the results from our screen and those obtained with TraSH pooled infection do not always agree. One complication of the TraSH analysis is that a certain degree of background noise will always be associated with high-throughput hybridization-based strategies. False positive and false negative results will occur (77, 89). It is also possible that differences in site of insertion played a role in disparate observations between our mutant analysis and TraSH results. If the insertion sites for the TraSH mutants were available, it would be interesting to compare the sites of insertion. Comparing sites of insertion could allow us to narrow down functional domains of a virulence factor. Screening pools of mutants, as was done in TraSH, may accentuate phenotypes by competition – strains that grow only marginally slower may be outcompeted by more fit strains upon group infections. This may explain why some ORFs predicted by TraSH as important in macrophages were not predicted by us to be important tested as a single mutant in isolation. Finally, it must be emphasized that for both TraSH and our screening Tn'blaTEM-1 mutants, follow up complementation experiments are required to prove any ORF encodes a virulence factor.

The library of mapped mutants generated in this study is a useful tool for exploring the roles that *M. tuberculosis* exported proteins play in pathogenesis and host-microbial interactions. In the future, this library can be further exploited to identify exported proteins required under different conditions. For example, the library could be screened for mutants unable to survive in IFN-gamma activated macrophages, where the engulfed bacilli are subjected to upregulated antimycobacterial defenses (84, 106). Mutants unable to survive this attack will help to identify exported proteins specifically

required to protect intracellular bacilli after the onset of adaptive immunity. The ability of this pathogen to block phagosome acidification and phagosome-lysosome fusion is a hallmark of *M. tuberculosis* virulence. Our library of mutants could also be screened for those unable to halt phagolysosomal fusion or acidification to find exported proteins required for the prevention of this macrophage process.

ACKNOWLEDGEMENTS

We would like to thank Drs. Bill Jacobs (Albert Eistein College of Medicine, Bronx, NY), Sunhee Lee (Duke University, Durham, NC) and Jordan Kriakov (Albert Einstein College of Medicine) for providing *Himar1* on pSH1 and members of the Braunstein lab for support. We would also like to thank Jonathan Sullivan, Erin McElvania-Tekippe and Marilia Oliveira for their assistance with animal experiments. This research was funded by the National Institute of Allergy and Infectious Disease of the NIH (AI054540 and AI070928 to M.B.). J.R.M was supported in part by an NIH Cell and Molecular Biology training grant (NIH 5-T32- GM008581) and a Dissertation Completion Fellowship from the Graduate School at UNC.

ATTRIBUTIONS

The experiments described here were conceived by MB and JRM. JAM assisted in cloning the Tn'blaTEM-1 transposon vector into the mycobacteriophage phAE159. All other experiments were conducted by JRM. JRM and MB wrote the manuscript.

REFERENCES

1. **Tuberculist website.** http://genolist.pasteur.fr/TubercuList/

2. **Abdallah, A. M., N. C. Gey van Pittius, P. A. Champion, J. Cox, J. Luirink, C. M. Vandenbroucke-Grauls, B. J. Appelmelk, and W. Bitter.** 2007. Type VII secretion--mycobacteria show the way. Nat Rev Microbiol **5:**883-91.

3. **Aires, J. R., and H. Nikaido.** 2005. Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD multidrug efflux transporter of *Escherichia coli.* J Bacteriol **187:**1923-9.

4. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J Mol Biol **215:**403-10.

5. **Andersen, P.** 1994. Effective vaccination of mice against *Mycobacterium tuberculosis* infection with a soluble mixture of secreted mycobacterial proteins. Infect Immun **62:**2536-44.

6. **Arruda, S., G. Bomfim, R. Knights, T. Huima-Byron, and L. W. Riley.** 1993. Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. Science **261:**1454-7.

7. **Azad, A. K., T. D. Sirakova, N. D. Fernandes, and P. E. Kolattukudy.** 1997. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J Biol Chem **272:**16741-5.

8. **Bardarov, S., J. Kriakov, C. Carriere, S. Yu, C. Vaamonde, R. A. McAdam, B. R. Bloom, G. F. Hatfull, and W. R. Jacobs, Jr.** 1997. Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **94:**10961-6.

9. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol **340:**783-95.

10. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

11. **Bengtsson, J., H. Tjalsma, C. Rivolta, and L. Hederstedt.** 1999. Subunit II of *Bacillus subtilis* cytochrome c oxidase is a lipoprotein. J Bacteriol **181:**685-8.

12. **Braibant, M., P. Gilot, and J. Content.** 2000. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. FEMS Microbiol Rev **24:**449-67.

13. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. *In* V. L. Clark and P. M. Bavoil (ed.), Meth Enzymol. **358**. Academic Press, London.

14. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-90.

15. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

16. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552'*phoA* in vitro transposition system. J Bacteriol **182:**2732-40.

17. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. Mol Microbiol **4:**1637-44.

18. **Camacho, L. R., P. Constant, C. Raynaud, M. A. Laneelle, J. A. Triccas, B. Gicquel, M. Daffe, and C. Guilhot.** 2001. Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. J Biol Chem **276:**19845-54.

19. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of *mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol Microbiol **34:**257-67.

20. **Camus, J. C., M. J. Pryor, C. Medigue, and S. T. Cole.** 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. Microbiol **148:**2967-73.

21. **Casali, N., and L. W. Riley.** 2007. A phylogenomic analysis of the Actinomycetales mce operons. BMC Genomics **8:**60.

22. **Casali, N., A. M. White, and L. W. Riley.** 2006. Regulation of the *Mycobacterium tuberculosis mce1* operon. J Bacteriol **188:**441-9.

23. **Charpentier, X., and E. Oswald.** 2004. Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. J Bacteriol **186:**5486-95.

24. **Chubb, A. J., Z. L. Woodman, F. M. da Silva Tatley, H. J. Hoffmann, R. R. Scholle, and M. R. Ehlers.** 1998. Identification of *Mycobacterium tuberculosis* signal sequences that direct the export of a leaderless beta-lactamase gene product in *Escherichia coli*. Microbiology **144:**1619-29.

25. **Claros, M. G., and G. von Heijne.** 1994. TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci **10:**685-6.

26. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-44.

27. **Converse, S. E., J. D. Mougous, M. D. Leavell, J. A. Leary, C. R. Bertozzi, and J. S. Cox.** 2003. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. Proc Natl Acad Sci U S A **100:**6121-6.

28. **Cox, J. S., B. Chen, M. McNeil, and W. R. Jacobs, Jr.** 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature **402:**79-83.

29. **Datta, N., and P. Kontomichalou.** 1965. Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae. Nature **208:**239-41.

30. **Davidson, A. L., E. Dassa, C. Orelle, and J. Chen.** 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev **72:**317-64, table of contents.

31. **de Chastellier, C., and L. Thilo.** 2006. Cholesterol depletion in *Mycobacterium avium*-infected macrophages overcomes the block in phagosome maturation and leads to the reversible sequestration of viable mycobacteria in phagolysosome-derived autophagic vacuoles. Cell Microbiol **8:**242-56.

32. **de Jong, M. F., Y. H. Sun, A. B. den Hartigh, J. M. van Dijl, and R. M. Tsolis.** 2008. Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the *Brucella* type IV secretion system. Mol Microbiol **70:**1378-96.

33. **De Rossi, E., P. Arrigo, M. Bellinzoni, P. A. Silva, C. Martin, J. A. Ainsa, P. Guglierame, and G. Riccardi.** 2002. The multidrug transporters belonging to major facilitator superfamily in *Mycobacterium tuberculosis*. Mol Med **8:**714-24.

34. **Domenech, P., M. B. Reed, and C. E. Barry, 3rd.** 2005. Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. Infect Immun **73:**3492-501.

35. **Domenech, P., M. B. Reed, C. S. Dowd, C. Manca, G. Kaplan, and C. E. Barry, 3rd.** 2004. The role of MmpL8 in sulfatide biogenesis and virulence of *Mycobacterium tuberculosis*. J Biol Chem **279:**21257-65.

36. **Downing, K. J., R. A. McAdam, and V. Mizrahi.** 1999. *Staphylococcus aureus* nuclease is a useful secretion reporter for mycobacteria. Gene **239:**293-299.

37. **Eicher, T., L. Brandstatter, and K. M. Pos.** 2009. Structural and functional aspects of the multidrug efflux pump AcrB. Biol Chem.

38. **El-Shazly, S., S. Ahmad, A. S. Mustafa, R. Al-Attiyah, and D. Krajci.** 2007. Internalization by HeLa cells of latex beads coated with mammalian cell entry (Mce) proteins encoded by the *mce3* operon of *Mycobacterium tuberculosis*. J Med Microbiol **56:**1145-51.

39. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev **61:**136-69.

40. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Characterization of novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* mutants hypersusceptible to beta-lactam antibiotics. J Bacteriol **187:**1892-900.

41. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. Microbiology **151:**521-32.

42. **Gatfield, J., and J. Pieters.** 2000. Essential role for cholesterol in entry of mycobacteria into macrophages. Science **288:**1647-50.

43. **Glickman, M. S., and W. R. Jacobs, Jr.** 2001. Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. Cell **104:**477-85.

44. **Guilhot, C. a. D., M.** 2008. Polyketides and Polyketide-Containing Glycolipids of *Mycobacterium tuberculosis*: Structure, Biosynthesis and Biological Activities. *In* S. H. a. R. Kaufmann, E. (ed.), Handbook of Tuberculosis. Wiley VCH, Weinheim, Germany.

45. **Hawley, R. S. W., Michelle Y.** 2003. Advanced Genetic Analysis: Finding Meaning in a Genome. Blackwell Science, Oxford, UK.

46. **Hofmann, K., and W. Stoffel.** 1993. TMbase - A database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler **374,:**166-170.

47. **Horwitz, M. A., B. W. Lee, B. J. Dillon, and G. Harth.** 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **92:**1530-4.

48. **Joshi, S. M., A. K. Pandey, N. Capite, S. M. Fortune, E. J. Rubin, and C. M. Sassetti.** 2006. Characterization of mycobacterial virulence genes through genetic interaction mapping. Proc Natl Acad Sci U S A **103:**11760-5.

49. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12:**1652-62.

50. **Kana, B. D., E. A. Weinstein, D. Avarbock, S. S. Dawes, H. Rubin, and V. Mizrahi.** 2001. Characterization of the *cydAB*-encoded cytochrome bd oxidase from *Mycobacterium smegmatis*. J Bacteriol **183:**7076-86.

51. **Kolattukudy, P. E., N. D. Fernandes, A. K. Azad, A. M. Fitzmaurice, and T. D. Sirakova.** 1997. Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. Mol Microbiol **24:**263-70.

52. **Kumar, A., M. Bose, and V. Brahmachari.** 2003. Analysis of expression profile of mammalian cell entry (*mce*) operons of *Mycobacterium tuberculosis*. Infect Immun **71:**6083-7.

53. **Kumar, A., A. Chandolia, U. Chaudhry, V. Brahmachari, and M. Bose.** 2005. Comparison of mammalian cell entry operons of mycobacteria: in silico analysis and expression profiling. FEMS Immunol Med Microbiol **43:**185-95.

54. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun.

55. **Lamichhane, G., S. Tyagi, and W. R. Bishai.** 2005. Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. Infect Immun **73:**2533-40.

56. **Lee, S., J. Kriakov, C. Vilcheze, Z. Dai, G. F. Hatfull, and W. R. Jacobs, Jr.** 2004. Bxz1, a new generalized transducing phage for mycobacteria. FEMS Microbiol Lett **241:**271-6.

57. **Lefevre, P., O. Denis, L. De Wit, A. Tanghe, P. Vandenbussche, J. Content, and K. Huygen.** 2000. Cloning of the gene encoding a 22-kilodalton cell surface antigen of *Mycobacterium bovis* BCG and analysis of its potential for DNA vaccination against tuberculosis. Infect Immun **68:**1040-7.

58. **Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi.** 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. J Bacteriol **177:**59-65.

59. **Machowski, E. E., R. A. McAdam, K. M. Derbyshire, and V. Mizrahi.** 2000. Construction and application of mycobacterial reporter transposons. Gene **253:**67-75.

60. **Malen, H., F. S. Berven, K. E. Fladmark, and H. G. Wiker.** 2007. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. Proteomics **7:**1702-18.

61. **Manoil, C., J. J. Mekalanos, and J. Beckwith.** 1990. Alkaline phosphatase fusions: sensors of subcellular location. J Bacteriol **172:**515-8.

62. **Marketon, M. M., R. W. DePaolo, K. L. DeBord, B. Jabri, and O. Schneewind.** 2005. Plague bacteria target immune cells during infection. Science **309:**1739-41.

63. **Marmiesse, M., P. Brodin, C. Buchrieser, C. Gutierrez, N. Simoes, V. Vincent, P. Glaser, S. T. Cole, and R. Brosch.** 2004. Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex. Microbiology **150:**483-96.

64. **Matsoso, L. G., B. D. Kana, P. K. Crellin, D. J. Lea-Smith, A. Pelosi, D. Powell, S. S. Dawes, H. Rubin, R. L. Coppel, and V. Mizrahi.** 2005. Function of the cytochrome bc1-aa3 branch of the respiratory network in mycobacteria and network adaptation occurring in response to its disruption. J Bacteriol **187:**6300-8.

65. **McCann, J. R., J. A. McDonough, M. S. Pavelka, and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. Microbiology **153:**3350-9.

66. **Megehee, J. A., J. P. Hosler, and M. D. Lundrigan.** 2006. Evidence for a cytochrome *bcc-aa3* interaction in the respiratory chain of *Mycobacterium smegmatis*. Microbiology **152:**823-9.

67. **Milano, A., M. R. Pasca, R. Provvedi, A. P. Lucarelli, G. Manina, A. L. Ribeiro, R. Manganelli, and G. Riccardi.** 2009. Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5-MmpL5 efflux system. Tuberculosis (Edinb) **89:**84-90.

68. **Mohn, W. W., R. van der Geize, G. R. Stewart, S. Okamoto, J. Liu, L. Dijkhuizen, and L. D. Eltis.** 2008. The actinobacterial *mce4* locus encodes a steroid transporter. J Biol Chem **283:**35368-74.

69. **Nakai, K., and P. Horton.** 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci **24:**34-6.

70. **Nikaido, H.** 2009. Multidrug resistance in bacteria. Annu Rev Biochem **78:**119-46.

71. **Onwueme, K. C., C. J. Vos, J. Zurita, J. A. Ferreras, and L. E. Quadri.** 2005. The dimycocerosate ester polyketide virulence factors of mycobacteria. Prog Lipid Res **44:**259-302.

72. **Pajon, R., D. Yero, A. Lage, A. Llanes, and C. J. Borroto.** 2006. Computational identification of beta-barrel outer-membrane proteins in *Mycobacterium tuberculosis* predicted proteomes as putative vaccine candidates. Tuberculosis (Edinb) **86:**290-302.

73. **Peyron, P., C. Bordier, E. N. N'Diaye, and I. Maridonneau-Parini.** 2000. Nonopsonic phagocytosis of *Mycobacterium kansasii* by human neutrophils depends on cholesterol and is mediated by CR3 associated with glycosylphosphatidylinositol-anchored proteins. J Immunol **165:**5186-91.

74. **Piddock, L. J.** 2006. Multidrug-resistance efflux pumps - not just for resistance. Nat Rev Microbiol **4:**629-36.

75. **Pluckthun, A., and J. R. Knowles.** 1987. The consequences of stepwise deletions from the signal-processing site of beta-lactamase. J Biol Chem **262:**3951-7.

76. **Prinz, W. A., and J. Beckwith.** 1994. Gene fusion analysis of membrane protein topology: a direct comparison of alkaline phosphatase and beta-lactamase fusions. J Bacteriol **176:**6410-3.

77. **Rengarajan, J., B. R. Bloom, and E. J. Rubin.** 2005. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A **102:**8327-32.

78. **Richter, O. M., and B. Ludwig.** 2003. Cytochrome c oxidase--structure, function, and physiology of a redox-driven molecular machine. Rev Physiol Biochem Pharmacol **147:**47-74.

79. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

80. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-36.

81. **Rousseau, C., N. Winter, E. Pivert, Y. Bordat, O. Neyrolles, P. Ave, M. Huerre, B. Gicquel, and M. Jackson.** 2004. Production of phthiocerol dimycocerosates protects *Mycobacterium tuberculosis* from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection. Cell Microbiol **6:**277-87.

82. **Rubin, E. J., B. J. Akerley, V. N. Novik, D. J. Lampe, R. N. Husson, and J. J. Mekalanos.** 1999. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. Proc Natl Acad Sci U S A **96:**1645-50.

83. **Russell, D. G.** 2001. *Mycobacterium tuberculosis*: here today, and here tomorrow. Nat Rev Mol Cell Biol **2:**569-77.

84. **Russell, D. G.** 2008. *Mycobacterium tuberculosis:* life and death in the phagosome, p. 307-322. *In* S. H. Kaufmann and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. Wiley-VCH.

85. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? Nat Rev Microbiol **5:**39-47.

86. **Saini, N. K., M. Sharma, A. Chandolia, R. Pasricha, V. Brahmachari, and M. Bose.** 2008. Characterization of Mce4A protein of *Mycobacterium tuberculosis*: role in invasion and survival. BMC Microbiol **8:**200.

87. **Sambrook, J., and D. W. Russell.** 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

88. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

89. **Sassetti, C. M., and E. J. Rubin.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-94.

90. **Shimono, N., L. Morici, N. Casali, S. Cantrell, B. Sidders, S. Ehrt, and L. W. Riley.** 2003. Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the *mce1* operon. Proc Natl Acad Sci U S A **100:**15918-23.

91. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:**175-82.

92. **Stewart, G. R., J. Patel, B. D. Robertson, A. Rae, and D. B. Young.** 2005. Mycobacterial mutants with defective control of phagosomal acidification. PLoS Pathog **1:**269-78.

93. **Sulzenbacher, G., S. Canaan, Y. Bordat, O. Neyrolles, G. Stadthagen, V. Roig-Zamboni, J. Rauzier, D. Maurin, F. Laval, M. Daffe, C. Cambillau, B. Gicquel, Y. Bourne, and M. Jackson.** 2006. LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. Embo J **25:**1436-44.

94. **Tadayyon, M., and J. K. Broome-Smith.** 1992. TnblaM: a transposon for directly tagging bacterial genes encoding cell envelope and secreted proteins. Gene **111:**21-6.

95. **Takeda, K., H. Miyatake, N. Yokota, S. Matsuyama, H. Tokuda, and K. Miki.** 2003. Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. Embo J **22:**3199-209.

96. **Takeda, K., H. Miyatake, N. Yokota, S. I. Matsuyama, H. Tokuda, and K. Miki.** 2003. Crystallization and preliminary crystallographic study of the outer-membrane lipoprotein receptor LolB, a member of the lipoprotein localization factors. Acta Crystallogr D Biol Crystallogr **59:**1224-6.

97. **Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos.** 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc Natl Acad Sci U S A **84:**2833-7.

98. **Tekaia, F., S. V. Gordon, T. Garnier, R. Brosch, B. G. Barrell, and S. T. Cole.** 1999. Analysis of the proteome of *Mycobacterium tuberculosis* in silico. Tuber Lung Dis **79:**329-42.

99. **Timm, J., M. G. Perilli, C. Duez, J. Trias, G. Orefici, L. Fattorini, G. Amicosante, A. Oratore, B. Joris, J. M. Frere, and et al.** 1994. Transcription and expression analysis, using *lacZ* and *phoA* gene fusions, of *Mycobacterium fortuitum* beta-lactamase genes cloned from a natural isolate and a high-level beta-lactamase producer. Mol Microbiol **12:**491-504.

100. **Timme, T. L., and P. J. Brennan.** 1984. Induction of bacteriophage from members of the *Mycobacterium avium, Mycobacterium intracellulare*, *Mycobacterium scrofulaceum* serocomplex. J Gen Microbiol **130:**2059-66.

101. **Tokuda, H., and S. Matsuyama.** 2004. Sorting of lipoproteins to the outer membrane in *E. coli*. Biochim Biophys Acta **1693:**5-13.

102. **Trivedi, O. A., P. Arora, V. Sridharan, R. Tickoo, D. Mohanty, and R. S. Gokhale.** 2004. Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. Nature **428:**441-5.

103. **Trivedi, O. A., P. Arora, A. Vats, M. Z. Ansari, R. Tickoo, V. Sridharan, D. Mohanty, and R. S. Gokhale.** 2005. Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. Mol Cell **17:**631-43.

104. **Tseng, T. T., K. S. Gratwick, J. Kollman, D. Park, D. H. Nies, A. Goffeau, and M. H. Saier, Jr.** 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. J Mol Microbiol Biotechnol **1:**107-25.

105. **Van der Geize, R., K. Yam, T. Heuser, M. H. Wilbrink, H. Hara, M. C. Anderton, E. Sim, L. Dijkhuizen, J. E. Davies, W. W. Mohn, and L. D. Eltis.** 2007. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. Proc Natl Acad Sci U S A **104:**1947-52.

106. **Vergne, I., J. Chua, S. B. Singh, and V. Deretic.** 2004. Cell biology of *mycobacterium tuberculosis* phagosome. Annu Rev Cell Dev Biol **20:**367-94.

107. **WorldHealthOrganization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.] http://www.who.int/mediacentre/factsheets/fs104/en/

108. **Xiong, Y., M. J. Chalmers, F. P. Gao, T. A. Cross, and A. G. Marshall.** 2005. Identification of *Mycobacterium tuberculosis* H37Rv integral membrane proteins by one-dimensional gel electrophoresis and liquid chromatography electrospray ionization tandem mass spectrometry. J Proteome Res **4:**855-61.

109. **Zapun, A., C. Contreras-Martel, and T. Vernet.** 2008. Penicillin-binding proteins and beta-lactam resistance. FEMS Microbiol Rev **32:**361-85.

Supplemental Table 4.1 Location of Tn'blaTEM-1 insertions resulting in carbenicillin resistance

Rv number	Gene	Tn insert site (AA/Total AA)	Export Signal[†]	Prior Proof of Export[#]
Rv0064		346/979	TM (8)	(1, 6)
Rv0064		347/979	TM (8)	(1, 6)
Rv0072**		199/349	TSP+TM(5)	
Rv0092	*ctpA*	204/761	TM (7)	
Rv0092	*ctpA*	204/761	TM (7)	
Rv0116c		90/251	SSP	(6)
Rv0125	*pepA*	228/355	SSP	(1, 6)
Rv0169*	*mce1A*	449/455	TM (2)	(2)
Rv0170*	*mce1B*	223/347	SSP	(6)
Rv0171*	*mce1C*	221/516	SSP	
Rv0171*	*mce1C*	512/516	SSP	
Rv0171*	*mce1C*	252/516	SSP	
Rv0172*	*mce1D*	67/531	SSP	(6)
Rv0173*	*lprK*	209/391	LSP	(6)
Rv0174*	*mce1F*	336/516	SSP	(6)
Rv0175*	*mce-assoc*	203/214	TM (1)	
Rv0175*	*mce-assoc*	203/214	TM (1)	
Rv0178*	*mce-assoc*	171/245	TM (1)	
Rv0178*	*mce-assoc*	171/245	TM (1)	
Rv0197	*betP*	375/594	TM (12)	
Rv0199*		74/220	TM (3)	
Rv0199*		74/220	TM (3)	
Rv0204c		52/413	TM (7)	
Rv0209*		133/362	TM (3)	
Rv0265c		42/331	TSP	(6)
Rv0312		492/621	TM (6)	
Rv0346c	*ansP2*	117/488	TM (5)	
Rv0361*		163/276	TM (1)	
Rv0402c*	*mmpL1*	816/959	TM (11)	(6)
Rv0402c*	*mmpL1*	816/959	TM (11)	(6)
Rv0412c		370/440	TM (4-5)	
Rv0418	*lpqL*	29/501	LSP	
Rv0432	*sodC*	106/241	LSP	
Rv0450c*	*mmpL4*	254/967	TM (12)	(1)
Rv0450c*	*mmpL4*	212/967	TM (12)	(1)
Rv0450c*	*mmpL4*	197/967	TM (12)	(1)
Rv0453*	*PPE11*	20/519	TM (6)	(6)
Rv0483	*lprQ*	445/451	TSP	
Rv0483	*lprQ*	445/451	TSP	
Rv0483	*lprQ*	445/451	TSP	
Rv0483	*lprQ*	445/451	TSP	
Rv0506*	*mmpS2*	29/148	TM (2)	(6)
Rv0506*	*mmpS2*	29/148	TM (2)	(6)
Rv0517*		234/503	TM (10)	
Rv0534c	*menA*	101/240	TM (7)	
Rv0583c*	*lpqN*	34/229	LSP	(6)
Rv0594*	*mce2F*	476/517	SSP	
Rv0676c*	*mmpL5*	457/965	TM (12)	(1)

Rv number	Gene	Tn insert site (AA/Total AA)	Export Signal[†]	Prior Proof of Export[#]
Rv0677c*	*mmpS5*	65/143	SSP	(6)
Rv0783c	*emrB*	530/541	TM (13)	
Rv0837c**		184/343	TM (1)	
Rv0876c		449/549	TM (10)	(6)
Rv0931c	*pknD*	592/665	TM (1)	
Rv0931c	*pknD*	583/665	TM (1)	
Rv0934	*pstSI*	268/375	LSP	(1, 6)
Rv0983	*pepD (mtb32b)*	236/465	TM (2)	
Rv1004c**		417/420	TM (7)	
Rv1004c**		417/420	TM (7)	
Rv1009	*rpfB*	93/363	LSP	
Rv1078**	*pra*	128/241	TM (3)	
Rv1096		288/292	SSP	
Rv1096		288/293	SSP	
Rv1157c*		173/372	SSP	
Rv1164	*narI*	156/247	TM (5)	
Rv1164	*narI*	156/247	TM (5)	
Rv1230c*		94/412	TM (4)	
Rv1230c*		94/412	TM (4)	
Rv1239c	*corA*	351/367	TM (2)	
Rv1319c		80/535	TM (6)	
Rv1319c		80/535	TM (6)	
Rv1319c		80/535	TM (6)	
Rv1368	*lprF*	205/262	LSP	
Rv1419**		31/158	SSP	(6)
Rv1435c**		97/234	SSP+TM (2)	
Rv1567c**		17/111	TM (2)	
Rv1567c**		17/111	TM (2)	
Rv1591*		79/322	TM (4)	
Rv1591*		79/322	TM (4)	
Rv1591*		79/322	TM (4)	
Rv1635c*		224/557	TM (10)	
Rv1648**		221/269	TM (8)	
Rv1707*		378/487	TM (11)	
Rv1707*		28/487	TM (11)	
Rv1728c*		68/257	TM (2)	
Rv1728c*		68/257	TM (2)	
Rv1728c*		68/257	TM (2)	
Rv1728c*		68/257	TM (2)	
Rv1743	*pknE*	393/567	TM (1)	
Rv1774c*	*TB8.4*	35/111	SSP	(6)
Rv1779c		140/598	TM (4)	
Rv1819c		240/640	TM (6)	
Rv1836c		507/678	TM (4)	
Rv1860*	*apa, mpt32*	197/326	SSP	(6)
Rv1886c*	*fbpB (Ag85B)*	274/326	SSP	(1, 6)
Rv1886c*	*fbpB (Ag85B)*	274/326	SSP	(1, 6)
Rv1886c*	*fbpB (Ag85B)*	201/326	SSP	(1, 6)
Rv1887*		294/380	TM (2)	(1, 6)
Rv1887*		294/380	TM (2)	(1, 6)
Rv1891*		70/136	SSP	
Rv1984c	*cfp21*	160/218	SSP	(6)

Rv number	Gene	Tn insert site (AA/Total AA)	Export Signal[†]	Prior Proof of Export[#]
Rv2040c		245/307	TM (7)	
Rv2068c^	*blaC*	16/308^	TSP	(6, 7)
Rv2080**	*lppJ*	167/188	LSP	(6)
Rv2091c*		121/245	TM (1)	
Rv2113*		333/398	TM (8)	
Rv2113*		333/398	TM (8)	
Rv2113*		333/398	TM (8)	
Rv2113*		333/398	TM (8)	
Rv2113*		333/398	TM (8)	
Rv2127	*ansP1*	446/490	TM (12)	
Rv2200c	*ctaC*	351/364	TM (3)	(6)
Rv2200c	*ctaC*	351/364	TM (3)	(6)
Rv2203*		88/231	TM (1)	
Rv2240c*		106/266	TM (3)	
Rv2240c*		106/266	TM (3)	
Rv2240c*		106/266	TM (3)	
Rv2240c*		106/266	TM (3)	
Rv2264c		517/593	SSP+TM (1)	
Rv2284	*lipM*	30/432	LSP	
Rv2284	*lipM*	236/432	LSP	
Rv2290**	*lppO*	55/172	LSP	(3)
Rv2290**	*lppO*	55/172	LSP	(3)
Rv2290**	*lppO*	55/172	LSP	(3)
Rv2290**	*lppO*	55/172	LSP	(3)
Rv2290**	*lppO*	55/172	LSP	(3)
Rv2301	*cfp25,cut2*	25/231	SSP	(6)
Rv2320c	*rocE*	384/477	SSP+TM (10)	
Rv2339*	*mmpL9*	409/963	TM (12)	
Rv2394	*ggtB*	576/644	TM (4)	
Rv2394	*ggtB*	576/644	TM (4)	
Rv2395		20/223	TM (16)	
Rv2443	*dctA*	124/492	TM (9)	
Rv2443	*dctA*	124/492	TM (9)	
Rv2585c		53/588	LSP	(6)
Rv2599**		110/144	SSP	
Rv2639c		40/111	TM (4)	
Rv2721c*		407/700	TM (2)	(6)
Rv2905*	*lppW*	263/315	LSP	(6)
Rv2905*	*lppW*	31/315	LSP	(6)
Rv2932	*ppsB*	614/1557	TM (6)	
Rv2938	*drrC*	67/277	TM (6)	
Rv2945c**	*lppX*	82/234	LSP	(1, 5, 6)
Rv3036c	*TB 22.2*	47/228	SSP	(6)
Rv3090**		64/296	SSP	
Rv3103c**		38/146	TM (1)	
Rv3209*		161/187	SSP	
Rv3209*		74/187	SSP	
Rv3209*		74/187	SSP	
Rv3209*		74/187	SSP	
Rv3253c		49/496	TM (12)	
Rv3253c		56/496	TM (12)	

	PPE56			
Rv3350c*		932/3717	TM (30-79)	
Rv number	**Gene**	**Tn insert site (AA/Total AA)**	**Export Signal[†]**	**Prior Proof of Export[#]**
Rv3267		458/499	SSP	(1, 6)
Rv3390*	*lpqD*	39/237	LSP	
Rv3390*	*lpqD*	149/237	LSP	
Rv3413c*		187/300	TM (1)	(6)
Rv3476c	*kgtP*	265/450	TM (11)	
Rv3484	*cpsA*	347/513	SSP+TM (1)	(6)
Rv3494	*mce4F*	539/565	TM (5)	
Rv3496c	*mce4D*	38/452	TM (1)	
Rv3496c	*mce4D*	38/452	TM (1)	
Rv3497c	*mce4C*	345/358	SSP+TM (1)	
Rv3497c	*mce4C*	260/358	SSP+TM (1)	
Rv3498c	*mce4B*	350/351	SSP	
Rv3584*	*lpqE*	143/183	LSP	(6)
Rv3689*		384/452	TM (6)	
Rv3779**		616/667	TM (13)	
Rv3802c		336/336	SSP	
Rv3802c		336/336	SSP	
Rv3802c		336/336	SSP	
Rv3802c		336/336	SSP	
Rv3802c		336/336	SSP	
Rv3802c		336/336	SSP	
Rv3835		385/449	TM (2)	(6)
Rv3835		385/449	TM (2)	(6)
Rv3861**		467/481	TM (1)	
Rv3901c**		57/150	SSP	
Rv3910		1075/1185	TM (15)	
Rv3910		1075/1185	TM (15)	

[†] As predicted by the bioinformatic algorithms identified in Materials and Methods. Abbreviations: AA, amino acid; SSP, Sec signal peptide; TSP, Tat signal peptide; LSP, lipoprotein signal peptide; TM (#), transmembrane domains [number predicted by TMHMM (8) and TMPred (4)].
*, Loci with at least 30% similarity at the amino acid level found only in mycobacteria
**, Loci with at least 30% similarity at the amino acid level found only in the virulent *Mycobacterium* species *M. avium, M. bovis, M. leprae, M. marinum, M. tuberculosis* or *M. ulcerans*
^, Tn'blaTEM-1 was inserted immediately downstream of a portion of the BlaC signal sequence that remains after deletion of the majority of the *blaC* coding region.
[#], References that describe export of respective ORF are numbered in parentheses.

REFERENCES

1. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552'*phoA* in vitro transposition system. J Bacteriol **182:**2732-40.

2. **Chitale, S., S. Ehrt, I. Kawamura, T. Fujimura, N. Shimono, N. Anand, S. Lu, L. Cohen-Gould, and L. W. Riley.** 2001. Recombinant *Mycobacterium tuberculosis* protein associated with mammalian cell entry. Cell Microbiol **3:**247-54.

3. **Gomez, M., S. Johnson, and M. L. Gennaro.** 2000. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach. Infect Immun **68:**2323-7.

4. **Hofmann, K., and W. Stoffel.** 1993. TMbase - A database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler **374,:**166-170.

5. **Lefevre, P., O. Denis, L. De Wit, A. Tanghe, P. Vandenbussche, J. Content, and K. Huygen.** 2000. Cloning of the gene encoding a 22-kilodalton cell surface antigen of *Mycobacterium bovis* BCG and analysis of its potential for DNA vaccination against tuberculosis. Infect Immun **68:**1040-7.

6. **Malen, H., F. S. Berven, K. E. Fladmark, and H. G. Wiker.** 2007. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. Proteomics **7:**1702-18.

7. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-79.

8. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:**175-82.

CHAPTER 5

DISCUSSION

Mycobacterium tuberculosis is the causative agent of tuberculosis, arguably the most successful bacterial disease in the history of mankind. Despite the fact that tuberculosis is a curable disease, an estimated 2 billion people are infected by *M. tuberculosis,* and the disease kills nearly 2 million people each year (56). To cause disease, the bacterium must survive and replicate in macrophages - a cellular environment normally hostile to invading bacteria. *M. tuberculosis* manipulates macrophage responses to prevent its own destruction. The bacilli also protect themselves against attack mounted by the innate immune response (42, 43, 54). As is true for other bacterial pathogens (17), *M. tuberculosis* exported proteins play central roles in promoting pathogen survival in the host.

In this thesis research the overarching goal was to identify exported proteins of *M. tuberculosis* and determine if they play roles in virulence. The unifying theme throughout this dissertation is the development and application of new genetic reporter systems to identify *M. tuberculosis* exported proteins. In Chapter 2, the primary objective was to identify the substrates of the conserved Tat export pathway of *M. tuberculosis*. Chapter 3 describes our efforts to create a genetic tool that could be used to report on *M. tuberculosis* proteins exported by intracellular bacilli. Finally, the focus of Chapter 4 was to identify *M. tuberculosis* exported proteins important for growth in

macrophages on a genome-wide scale. In this Discussion, I briefly review the results of this thesis research in the context of *M. tuberculosis* research as a whole. I will also discuss potential future experiments that will take advantage of these new genetic tools.

Identification of functional Tat signal sequences in *M. tuberculosis* proteins

As discussed in Chapter 2, our first major objective was to identify substrates of the Twin-Arginine Translocation (Tat) pathway in *M. tuberculosis*. The Tat pathway has an important role in virulence for other bacterial pathogens (7, 9, 15, 22, 40) . Mycobacteria possess a functional Tat pathway, and the *M. tuberculosis* β-lactamase BlaC is a Tat substrate (28, 32). Further, phospholipase C (Plc) enzymes, which are known *M. tuberculosis* virulence factors, have functional Tat signal sequences in their N-termini (28, 29, 33). Therefore, evidence exists that the Tat pathway in *M. tuberculosis* is also important for virulence. We hypothesized that the Tat pathway exports additional proteins with roles in pathogenesis and sought to identify these proteins. It would be ideal to investigate the role of the Tat pathway by analyzing the phenotypes of *M. tuberculosis* Δ*tat* deletion strains. However, the Tat pathway is essential (44) in *M. tuberculosis,* which makes using Δ*tat* mutants to investigate the effect of the Tat pathway on virulence not an option. It should be noted that the essentiality of the Tat pathway makes its components attractive targets for new anti-tuberculosis therapies.

We set out to identify proteins exported by the Tat pathway in *M. tuberculosis* using an alternative genetic approach. Using ‘BlaC, as a reporter specific for Tat exported substrates, we selected ORFs that possess functional Tat signal sequences from a *M. tuberculosis* genomic library. In this work, exported fusions were identified by

selecting β-lactam resistant colonies following electroporation of the *M. tuberculosis*-'*blaC* library into *M. smegmatis*. This is the first time any Tat-specific reporter was used in a positive selection strategy to identify proteins with Tat signal sequences. Prior to these studies, there were only predictions of potential *M. tuberculosis* Tat substrates based on bioinformatic algorithms (2, 31, 38). These algorithms are trained using known and predicted Tat substrates from a limited number of bacterial species, which notably does not include mycobacteria. In comparison, our approach was an unbiased way of identifying *M. tuberculosis* sequences able to export a Tat specific reporter.

We identified 13 *M. tuberculosis* signal sequences capable of exporting 'BlaC in a Tat specific manner. The above mentioned Tat prediction algorithms identified 95 potential Tat substrates in *M. tuberculosis*. However, only 10 unique Tat signal sequences were identified in our library. The additional three were demonstrated as able to export 'BlaC by directly fusing the Tat export signal to 'BlaC and testing for β-lactam resistance on plates when the fusion is expressed in *M. smegmatis*. One of these 13 proteins was not predicted by any algorithm to be a Tat substrate, indicating that the unbiased library approach was useful in identifying new *M. tuberculosis* Tat substrates. Further, our 'BlaC reporter was recently used by another laboratory to demonstrate that four more potential *M. tuberculosis* proteins possess functional Tat signal sequences (26).

There was little overlap in the number of Tat-exported proteins predicted by three different algorithms. Experimentally evaluating potential Tat substrates is therefore critical, and will help to hone the criteria these algorithms use to identify Tat exported proteins from primary sequence. An unanticipated pitfall of this strategy was that 'BlaC was only exported and active when fused directly downstream of the cleavage site of Tat

signal sequences. This requirement likely reflects the fact that Tat substrates must be folded in the cytoplasm before they are exported (14). While the algorithms may overestimate the number of Tat substrates that are actually found in *M. tuberculosis*, the requirement for Tat signal sequences to be positioned immediately adjacent to the 'BlaC mature protein domain is at least partly responsible for the small number of potential Tat substrates we identified. Extended mature sequence of a foreign protein may interfere with proper folding of 'BlaC, and cause the protein to be degraded before it could be exported. Another possibility is that improperly folded hybrids may have been exported as inactive fusions. Thus, even though our goal was to comprehensively identify *M. tuberuclosis* proteins exported by the Tat pathway, there are likely more substrates encoded in the genome that we missed with our strategy. It may be worthwhile to revisit the library design and optimize it so that smaller genomic fragments, likely to contain a signal sequence only, would be inserted downstream of a promoter and upstream of '*blaC*. In this manner, we may be able to experimentally identify more Tat substrates of *M. tuberculosis*. The addition of a sequence encoding a flexible amino acid linker upstream of '*blaC* in the library vector may also allow for a wider range of fusions to drive export of the reporter in a Tat dependent manner. A linker between fused sequence and 'BlaC may allow the reporter to fold independently of the Tat sequence driving its export, increasing the number of detectable exported fusions.

Besides BlaC itself and the Plc enzymes, the substrates identified have amino acid sequence similarity to redox proteins, substrate-binding proteins involved in nutrient uptake and lipases. Aside from Plc, we have not yet shown that the respective full-length proteins have roles in pathogenesis or are exported exclusively through the Tat pathway

in *M. tuberculosis*. However, two of the substrates are predicted to be essential for in vitro growth of the bacilli, providing some clues as to why the Tat pathway components cannot be deleted in *M. tuberculosis* (44, 45). Further characterization of each substrate will be required to determine their dependence on the Tat pathway for export as well as their potential roles in virulence.

β-lactamase can function as a reporter of bacterial protein export during *M. tuberculosis* infection of host cells

In Chapter 3 of this dissertation, I describe the development of a β-lactamase reporter for general protein export that can be used when fused to substrates of Sec and Tat export pathways in mycobacteria. Established reporters for studying export of *M. tuberculosis* proteins, such as alkaline phosphatase (5, 23) and staphylococcus nuclease (16) have drawbacks; they report on protein export based on labor intensive colony phenotype screens, and can not be used directly in *M. tuberculosis* due to endogenous activity in the pathogen. Our β-lactamase reporter is derived from standard cloning vectors and has been successfully used in the past to report on the subcellular localization of proteins exported by both conserved and specialized pathways in bacteria (8, 10, 12, 25). We developed a system that allows us to identify colonies expressing *M. tuberculosis* reporter fusions using a selection, not a screen. We obtained the β-lactam sensitive Δ*blaC M. tuberculosis* mutant from a collaborator (18), and used it prove that a heterologous β-lactamase lacking its own signal sequence, '*blaTEM-1*, could be expressed and exported by mycobacterial Tat and Sec signal sequences directly in *M.*

tuberculosis. Further, unlike fusions to ‘BlaC, fusion of mature domains from foreign proteins fused to ‘BlaTEM-1 did not interfere with export or activity of the reporter.

Strategies previously used to identify exported proteins of *M. tuberculosis* have always been conducted with in vitro grown bacilli (5, 16, 39, 49). This is true for the library approach described in Chapter 2 of this dissertation, as well as for the transposon reporter approach described in Chapter 4. We hypothesize that there is a set of proteins only expressed or exported inside host cells and that these proteins, which are the best candidates for being virulence factors, are overlooked by evaluation of *in vitro* grown bacteria. Even for more genetically tractable intracellular bacterial pathogens, this is a challenging category of virulence factors to identify. Importantly, we demonstrate in Chapter 3 that when *M. tuberculosis* is growing in host cells, ‘BlaTEM-1 still works to report on export. Intracellular bacilli expressing a signal sequence-‘BlaTEM-1 fusion are protected from β-lactam treatment (27). In the future, a potential use for this reporter is to identify those proteins exclusively exported in macrophages. The strategy outlined below is one way we could work to achieve this goal. A library of *M. tuberculosis* genomic fragments fused to the ‘*blaTEM-1* reporter will be constructed. The resulting plasmid library will be electroporated into Δ*blaC M. tuberculosis*, and transformants would be pooled and used to infect macrophages or mice. Following uptake of the bacilli, infected macrophages or animals will be treated with β-lactam antibiotic. Intracellular *M. tuberculosis* not expressing an exported fusion should be sensitive to β-lactam treatment and fail to grow. However, strains in the library pool exporting a ‘BlaTEM-1 fusion will grow and become enriched during the intracellular infection. Surviving strains would then be tested for their ability to survive on plates in the presence

of β-lactam. The fusions of interest would be expressed and exported in vivo, and would therefore allow for survival during intracellular β-lactam treatment but not during in vitro β-lactam exposure. Following rounds of infection where strains surviving infection in β-lactam treated macrophages are used to re-infect β-lactam treated macrophages or mice, we should selectively enrich for those strains exporting fusions within macrophages. The fusions that are recovered should identify new proteins exclusively exported in vivo. Such proteins are strong candidates for being virulence factors. The Tat-specific ‘BlaC reporter could also be used in a similar assay to find Tat-dependent proteins only exported *in vivo*. However, because the ‘BlaTEM-1 reporter works with protein fusions that drive export via both the Tat and Sec pathways of mycobacteria, a second Tat-specific reporter might not be required. If this strategy is successful, it may be widely applicable to the identification of in vivo exported proteins in other intracellular pathogens.

Future use of the β-lactamase reporter might also include subcellular localization of *M. tuberculosis* exported proteins within infected macrophages. We could use commercially available anti-‘BlaTEM-1 antibodies in an immunoelectron microscopy assay to localize intracellular *M. tuberculosis* exported protein fusions to the reporter. Another possibility is to use the reporter in conjunction with an intracellular β-lactamase substrate that becomes fluorescent when cleaved. The substrate, CCF2 (57), has been used in the past to show translocation of Type III and type IV secreted bacterial effector ‘BlaTEM-1 fusions into host cells (10, 12, 25). Upon translocation of an active bacterial protein-‘BlaTEM-1 fusion into host cytoplasm, the CCF2 is cleaved and becomes fluorescent. The fluorescence is then detected by microscopy or plate reader. There are at

least three examples of *M. tuberculosis* proteins exported out from the bacilli to the phagosomal membrane and beyond. The translocation of these proteins from within the intracellular bacterium to the host phagosomal membrane or cytoplasm is detected by immunoelectron microscopy (3, 13, 55). Using a fluorescent substrate, it could one-day be possible to screen a library for *M. tuberculosis* 'BlaTEM-1 fusions exported into the cytoplasm of infected macrophages. Exported proteins of interest could also be directly fused to the 'BlaTEM-1 reporter and their intracellular location examined. Some modifications to the CCF2 system are required before it can be used to directly detect *M. tuberculosis* proteins exported beyond the phagosome. At this point in time, the CCF2 substrate appears to permeate the entire host cell, including the phagosome. Thus, all exported *M. tuberculosis* exported proteins fused to the reporter are likely to cleave the substrate, including those that never transit beyond the phagosomal membrane. Live *M. tuberculosis* must be manipulated under biosafety level 3 (BL3) conditions. Therefore, the analysis requires a fluorescence microscope in the BL3 laboratory. Alternatively, the CCF2 detection protocol must be made compatible with the fixation conditions used to kill intracellular *M. tuberculosis* before it is taken out of BL3 containment.

Development of a reporter transposon to identify exported virulence factors in *M. tuberculosis*

Finally, we constructed a β-lactamase reporter transposon for identifying exported proteins encoded throughout the genome of *M. tuberculosis*. The use of a reporter transposon is a powerful and classic strategy, but existing reporter transposons are not useful in *M. tuberculosis* due to endogenous activity. We built a new genetic tool to

allow this type of investigation to be conducted directly in the pathogenic bacillus. Many exported proteins are only predicted as such on the basis of bioinformatic algorithms. Further, much of the *M. tuberculosis* genome has no predicted function, and proof of export will help to define function for these proteins. Even for those well-established exported *M. tuberculosis* proteins, there are very few with directly demonstrated functions in pathogenesis.

We placed '*blaTEM-1* in a Himar-1 transposon. We chose this transposon backbone because it has been successfully used in other saturating mutagenesis strategies in *M. tuberculosis*. Transposon insertions that placed '*blaTEM-1* downstream and in frame with an export signal in the *M. tuberculosis* genome resulted in production of a protein- β-lactamase fusion. Export of these fusions led to β-lactam resistance on plates. Using the Tn'blaTEM-1 reporter transposon we built, we could positively select for in-frame insertions that lead to the export of *M. tuberculosis*-'BlaTEM-1 fusion proteins.

Our goal was to comprehensively identify the exported proteins encoded by the *M. tuberculosis* genome. It is predicted that 15-20% of the proteins encoded by *M. tuberculosis* are exported (11). This would equate to approximately 800 proteins. If 20% of these proteins are essential (21, 45), we predict there should be 640 proteins potentially identifiable with Tn'blaTEM-1. Here, we identified 118 unique exported proteins. However, a Poisson calculation predicts that even after selecting from over 81,000 transposon mutants, there is still a 20% chance of identifying new exported ORFs if we continued this project (20). We also predict that there is a set of exported proteins incapable of transporting active 'BlaTEM-1. We had hypothesized that the 'BlaTEM-1 reporter would also be compatible with substrates of mycobacterial specialized export

pathways such as ESX-1 (1) and SecA2 (37). However, we did not identify any of these proteins with our Tn'blaTEM-1 selection strategy. In fact, every protein identified with the reporter appears to be a classical exported protein in either possessing an N-terminal Sec or Tat-type signal sequence or TM domain.

We went on to directly test the feasibility of using 'BlaTEM-1 to report on translocation of unconventional exported proteins. We directly fused ESX-1 substrate ESAT-6 (19, 50) to either the N- or C- termini of 'BlaTEM-1 and were unable to detect export. ESAT-6 is known to form tight heterodimers with CFP-10, another small protein exported by the ESX-1 pathway (6, 30, 35, 36, 48, 50). These dimers form prior to and are likely required for export of both binding partners. The fusion to 'BlaTEM-1 fusion coudl have blocked this dimerization. The 'BlaTEM-1 reporter also did not work when fused to SodA, a SecA2 exported protein (4). Thus, these specialized export mechanisms are likely to be incompatible with the 'BlaTEM-1 reporter. To our knowledge, no reporter fusion has been successfully used to identify the subcellular location of any specialized secretion pathway substrate of *M. tuberculosis*.

Using our reporter transposon, we identified 118 exported proteins of *M. tuberculosis*. An attribute of our strategy is that the reporter is carried on a transposon. Thus, while identifying exported proteins, we simultaneously constructed a mapped library of *M. tuberculosis* transposon mutants lacking individual exported proteins. We screened our library for mutants for growth defects in resting macrophages, and identified eight attenuated mutants.

Four of the eight attenuated mutants, *mce2F::tn'blaTEM-1, lppX::tn'blaTEM-1, ppsB::tn'blaTEM-1* and *ctaC::tn'blaTEM-1*, had insertions in ORFs never before shown

to be important for growth in macrophages. For the other four mutants, *mce1A::tn'blaTEM-1, mce1B::tn'blaTEM-1, rv0199::tn'blaTEM-1* and *mmpS5::tn'blaTEM-1*, the ORFs receiving insertions are predicted to be important for growth in macrophages in large-scale pooled infections (34, 46, 51). None of these mutants had previously been tested individually for the ability to grow in macrophages. Three of the eight respective ORFs have a predicted function, the other five do not. Further characterization of the eight attenuated mutants we identified must be prioritized. It would be interesting to first focus on Rv0199. It is restricted to mycobacteria, and is in a genetic locus that is predicted to be required for intracellular and in vivo growth (46, 51). Genetic complementation and aerosol infection of mice are two assays that could be done immediately to better characterize the role of Rv0199 in pathogenesis.

We can further analyze our library of mutants in ways that can help determine the function of *M. tuberculosis* exported proteins in virulence. When resting macrophages engulf *M. tuberculosis*, the bacillus prevents acidification of the phagosomes in which they reside (54). The ability to block phagosome acidification is believed to be important to M. *tuberculosis* growth in macrophages. Activation of macrophages in culture by addition of IFNγ results in phagosome-lysosome fusion, acidification of the *M. tuberculosis* containing phagosome, and inhibition of intracellular *M. tuberculosis* growth (24, 47, 52, 53). Recently, an *M. tuberculosis* pH sensitive mutant was tested for its ability to grow in activated macrophages. This *M. tuberculosis* mutant is killed in activated macrophages while WT *M. tuberculosis* persists (52). Along these lines, our Tn'blaTEM-1 mutant library could next be screened for strains that are defective for survival in activated macrophages. The ability of *M. tuberculosis* to survive acidic

intracellular conditions is considered an important virulence mechanism for the bacillus, and this set of proteins may make attractive drug targets. Another virulence trait of *M. tuberculosis* is its ability to persist long after the onset of adaptive immunity following infection (41). This type screen may also help identify exported proteins important for *M. tuberculosis* persistence.

At the start of this work, we were interested in exported proteins because of their potential role in virulence, but had no genetic tools that could be used to identify and inactivate them directly in *M. tuberculosis*. In this thesis, we demonstrate that β-lactamase reporters are tools we previously lacked for characterization of the suite of *M. tuberculosis* exported proteins. Along with providing valuable information on the exported nature of 118 *M. tuberculosis* proteins, the reporter systems in themselves provide a valuable resource to the *M. tubercuosis* research community, where useful genetic tools are comparatively few. Future studies will expand on this work by exploiting the reporters and the Tn'blaTEM-l library to identify proteins exported by the bacilli in host cells, as well as those required for intracellular defenses against macrophage attack. A better understanding of *M. tuberculosis* exported proteins will ultimately reveal the ways in which this pathogen interacts with the host to cause disease.

REFERENCES

1. **Abdallah, A. M., N. C. Gey van Pittius, P. A. Champion, J. Cox, J. Luirink, C. M. Vandenbroucke-Grauls, B. J. Appelmelk, and W. Bitter.** 2007. Type VII secretion--mycobacteria show the way. Nat Rev Microbiol **5:**883-91.

2. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

3. **Berthet, F. X., J. Rauzier, E. M. Lim, W. Philipp, B. Gicquel, and D. Portnoi.** 1995. Characterization of the *Mycobacterium tuberculosis erp* gene encoding a potential cell surface protein with repetitive structures. Microbiology **141:**2123-30.

4. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

5. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552'phoA in vitro transposition system. J Bacteriol **182:**2732-40.

6. **Brodin, P., M. I. de Jonge, L. Majlessi, C. Leclerc, M. Nilges, S. T. Cole, and R. Brosch.** 2005. Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis*, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. J Biol Chem **280:**33953-9.

7. **Bronstein, P. A., M. Marrichi, S. Cartinhour, D. J. Schneider, and M. P. DeLisa.** 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae* pv. tomato DC3000 and its contribution to pathogenicity and fitness. J Bacteriol **187:**8450-61.

8. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. Mol Microbiol **4:**1637-44.

9. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. Mol Plant Microbe Interact **19:**200-12.

10. **Charpentier, X., and E. Oswald.** 2004. Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. J Bacteriol **186:**5486-95.

11. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-44.

12. **de Jong, M. F., Y. H. Sun, A. B. den Hartigh, J. M. van Dijl, and R. M. Tsolis.** 2008. Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the *Brucella* type IV secretion system. Mol Microbiol **70:**1378-96.

13. **Deghmane, A. E., H. Soualhine, H. Bach, K. Sendide, S. Itoh, A. Tam, S. Noubir, A. Talal, R. Lo, S. Toyoshima, Y. Av-Gay, and Z. Hmama.** 2007. Lipoamide dehydrogenase mediates retention of coronin-1 on BCG vacuoles, leading to arrest in phagosome maturation. J Cell Sci **120:**2796-806.

14. **DeLisa, M. P., P. Samuelson, T. Palmer, and G. Georgiou.** 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. J Biol Chem **277:**29825-31.

15. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. J Bacteriol **185:**760-71.

16. **Downing, K. J., R. A. McAdam, and V. Mizrahi.** 1999. *Staphylococcus aureus* nuclease is a useful secretion reporter for mycobacteria. Gene **239:**293-299.

17. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev **61:**136-69.

18. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. Microbiology **151:**521-32.

19. **Guinn, K. M., M. J. Hickey, S. K. Mathur, K. L. Zakel, J. E. Grotzke, D. M. Lewinsohn, S. Smith, and D. R. Sherman.** 2004. Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. Mol Microbiol **51:**359-70.

20. **Hawley, R. S. W., Michelle Y.** 2003. Advanced Genetic Analysis: Finding Meaning in a Genome. Blackwell Science, Oxford, UK.

21. **Lamichhane, G., M. Zignol, N. J. Blades, D. E. Geiman, A. Dougherty, J. Grosset, K. W. Broman, and W. R. Bishai.** 2003. A postgenomic method for

predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **100:**7213-8.

22. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. Infect Immun **74:**1768-76.

23. **Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi.** 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. J Bacteriol **177:**59-65.

24. **MacMicking, J. D., G. A. Taylor, and J. D. McKinney.** 2003. Immune control of tuberculosis by IFN-gamma-inducible LRG-47. Science **302:**654-9.

25. **Marketon, M. M., R. W. DePaolo, K. L. DeBord, B. Jabri, and O. Schneewind.** 2005. Plague bacteria target immune cells during infection. Science **309:**1739-41.

26. **Marrichi, M., L. Camacho, D. G. Russell, and M. P. DeLisa.** 2008. Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. J Biol Chem **283:**35223-35.

27. **McCann, J. R., J. A. McDonough, M. S. Pavelka, and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. Microbiology **153:**3350-9.

28. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-79.

29. **McDonough, J. A., J. R. McCann, E. M. Tekippe, J. S. Silverman, N. W. Rigel, and M. Braunstein.** 2008. Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol **190:**6428-38.

30. **Okkels, L. M., and P. Andersen.** 2004. Protein-protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis*. J Bacteriol **186:**2487-91.

31. **Pohlschroder, M., W. A. Prinz, E. Hartmann, and J. Beckwith.** 1997. Protein translocation in the three domains of life: variations on a theme. Cell **91:**563-6.

32. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-40.

33. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-17.

34. **Rengarajan, J., B. R. Bloom, and E. J. Rubin.** 2005. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A **102:**8327-32.

35. **Renshaw, P. S., K. L. Lightbody, V. Veverka, F. W. Muskett, G. Kelly, T. A. Frenkiel, S. V. Gordon, R. G. Hewinson, B. Burke, J. Norman, R. A. Williamson, and M. D. Carr.** 2005. Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. Embo J **24:**2491-8.

36. **Renshaw, P. S., P. Panagiotidou, A. Whelan, S. V. Gordon, R. G. Hewinson, R. A. Williamson, and M. D. Carr.** 2002. Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6*CFP-10 complex. Implications for pathogenesis and virulence. J Biol Chem **277:**21598-603.

37. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

38. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-50.

39. **Rosenkrands, I., A. King, K. Weldingh, M. Moniatte, E. Moertz, and P. Andersen.** 2000. Towards the proteome of *Mycobacterium tuberculosis.* Electrophoresis **21:**3740-56.

40. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. Infect Immun **73:**2020-32.

41. **Russell, D. G.** 2001. *Mycobacterium tuberculosis*: here today, and here tomorrow. Nat Rev Mol Cell Biol **2:**569-77.

42. **Russell, D. G.** 2008. *Mycobacterium tuberculosis:* life and death in the phagosome, p. 307-322. *In* S. H. Kaufmann and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. Wiley-VCH.

43. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? Nat Rev Microbiol **5:**39-47.

44. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-79.

45. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

46. **Sassetti, C. M., and E. J. Rubin.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-94.

47. **Schaible, U. E., S. Sturgill-Koszycki, P. H. Schlesinger, and D. G. Russell.** 1998. Cytokine activation leads to acidification and increases maturation of *Mycobacterium avium*-containing phagosomes in murine macrophages. J Immunol **160:**1290-6.

48. **Singh, A., D. Mai, A. Kumar, and A. J. Steyn.** 2006. Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. Proc Natl Acad Sci U S A **103:**11346-51.

49. **Sonnenberg, M. G., and J. T. Belisle.** 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. Infect Immun **65:**4515-24.

50. **Stanley, S. A., S. Raghavan, W. W. Hwang, and J. S. Cox.** 2003. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. Proc Natl Acad Sci U S A **100:**13001-6.

51. **Stewart, G. R., J. Patel, B. D. Robertson, A. Rae, and D. B. Young.** 2005. Mycobacterial mutants with defective control of phagosomal acidification. PLoS Pathog **1:**269-78.

52. **Vandal, O. H., L. M. Pierini, D. Schnappinger, C. F. Nathan, and S. Ehrt.** 2008. A membrane protein preserves intrabacterial pH in intraphagosomal *Mycobacterium tuberculosis*. Nat Med **14:**849-54.

53. **Vandal, O. H., J. A. Roberts, T. Odaira, D. Schnappinger, C. F. Nathan, and S. Ehrt.** 2009. Acid-susceptible mutants of *Mycobacterium tuberculosis* share hypersusceptibility to cell wall and oxidative stress and to the host environment. J Bacteriol **191:**625-31.

54. **Vergne, I., J. Chua, S. B. Singh, and V. Deretic.** 2004. Cell biology of *Mycobacterium tuberculosis* phagosome. Annu Rev Cell Dev Biol **20:**367-94.

55. **Walburger, A., A. Koul, G. Ferrari, L. Nguyen, C. Prescianotto-Baschong, K. Huygen, B. Klebl, C. Thompson, G. Bacher, and J. Pieters.** 2004. Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. Science **304:**1800-4.

56. **WorldHealthOrganization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.] http://www.who.int/mediacentre/factsheets/fs104/en/

57. **Zlokarnik, G., P. A. Negulescu, T. E. Knapp, L. Mere, N. Burres, L. Feng, M. Whitney, K. Roemer, and R. Y. Tsien.** 1998. Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter. Science **279:**84-8.